**RESEARCH STRATEGY**

**SIGNIFICANCE**

A recent WHO report stated “Antimicrobial resistance (AMR) within a wide range of infectious agents is a growing public health threat of broad concern to countries and multiple sectors. Increasingly, governments around the world are beginning to pay attention to a problem so serious that it threatens the achievements of modern medicine. A post-antibiotic era—in which common infections and minor injuries can kill—far from being an apocalyptic fantasy, is instead a very real possibility for the 21st century.” ([1](#_ENREF_1))

Pathogenic microbes are acquiring resistance to antibiotics faster than we can develop new compounds. A new paradigm for fighting these multi drug resistant pathogens is required. We propose to use a novel affinity selection technology to attack the virulence factors of these organisms by evolving protein ligands that will bind to, and inactivate key virulence structures (enzymes, binding domains, and toxins), rendering the bacterium harmless. Our approach consists of attacking an under-exploited pathogen vulnerability, that is, their virulence factors, by creating high affinity protein-based inhibitors against their virulence factors. The concept is to generate high-affinity ligands to the target virulence factors some of which we ***hypothesize*** will act as inhibitors. The targets in the R21 phase are two of the major virulence factors produced by *Streptococcus pneumoniae* (CbpA & Pln) listed as a Serious Health Threat by the CDC ([2](#_ENREF_2)). In the R33 phase we will also target two toxins of *Clostridium difficile* (the cell receptor binding modules of toxins A & B to which inactivating antibodies bind) listed by the CDC as an Urgent Health Threat ([2](#_ENREF_2))

Affinity selection (e.g. candidate ligand binding to a virulence factor) is the central technology used to identify new protein-based therapeutics. Phage display, the current state of the art for affinity selection, while effective is a cumbersome manual technology allowing at best one or two generations of mutagenesis and selection per day. Moreover, current affinity selection systems do not always produce high affinities or the types of variants ([3](#_ENREF_3)) desired.

Inspired by a system for continuous evolution ([4-6](#_ENREF_4)) and enabled by a recent advance in protein design by evolution called PACE ([7](#_ENREF_7)) Innatrix has designed and is testing a novel system, that addresses the features missing in those two systems that prevent easy application to the generation of new high affinity protein ligands. Innatrix will implement a novel, fully automated affinity selection system based on evolution of protein-protein binding we call Phage Assisted Three-Hybrid Evolution (PATHE). This is a new affinity selection technology with far reaching capabilities, and represents a radical new method of linking evolutionary fitness to protein-protein interaction. PATHE is versatile and can be employed for the development of an initial ligand, further optimization of an existing ligand, or the reoptimization of a ligand to overcome resistance. An additional benefit of this project is the value of the proposed PATHE technology itself.

***CDC Report Information on Streptococcus pneumoniae*** ([2](#_ENREF_2)) “*Streptococcus pneumoniae* (*S. pneumoniae*, or *pneumococcus*) is the leading cause of bacterial pneumonia and meningitis in the United States. It also is a major cause of bloodstream infections and ear and sinus infections.

Resistance of Concern

*S. pneumoniae* has developed resistance to drugs in the penicillin and erythromycin groups. Examples of these drugs include amoxicillin and azithromycin (Zithromax, Z-Pak). *S. pneumoniae* has also developed resistance to less commonly used drugs.

Public Health Threat

Pneumococcal disease, whether or not resistant to antibiotics, is a major public health problem. Pneumococcal disease causes 4 million disease episodes and 22,000 deaths annually. Pneumococcal ear infections (otitis media) are the most common type of pneumococcal disease among children, causing 1.5 million infections that often result in antibiotic use. Pneumococcal pneumonia is another important form of pneumococcal disease. Each year, nearly 160,000 children younger than 5 years old see a doctor or are admitted to the hospital with pneumococcal pneumonia. Among adults, over 600,000 seek care for or are hospitalized with pneumococcal pneumonia. Pneumococcal pneumonia accounts for 72% of all direct medical costs for treatment of pneumococcal disease. In 30% of severe *S. pneumoniae* cases, the bacteria are fully resistant to one or more clinically relevant antibiotics. Resistant infections complicate treatment and can result in almost 1,200,000 illnesses and 7,000 deaths per year. Cases of resistant pneumococcal pneumonia result in about 32,000 additional doctor visits and about 19,000 additional hospitalizations each year. The excess costs associated with these cases are approximately $96 million.” ([2](#_ENREF_2))

Thus, there is compelling need to create a new mode of attack on pathogens. Our proposed affinity selection system supports such a new mode that can be easily used by a broad range of investigators to construct new protein-based therapeutics, biosensors, and diagnostics. ***The overall goal of this project is to use the PATHE system to create protein-based therapeutics that can be used to inhibit the activity of bacterial pathogens regardless of the drug-resistant status of the pathogens.***

**Novel concept and innovation:** A system for continuous laboratory evolution (mutagenesis and selection) selecting for proteins to bind to a target protein of interest.

**INNOVATION**

Engineering proteins is one of the foundations of biotechnology, but the process of multiple rounds of design, mutagenesis, and identification of improved function is arduous and failure prone. Biology, in the form of continuous natural selection, greatly outstrips the human experience in inventing and optimizing protein function. Occasionally, the power of continuous selection has been incorporated into an experimental system to artificially optimize a protein function. However, these systems are exceptionally difficult to adapt to each new purpose, and have not been widely applied. As with any disruptive technology, the revolution will start when the technical blocks to widespread application are eliminated. We address the key bottlenecks that have prevented widespread application of the PACE system which implements continuous laboratory evolution ([7](#_ENREF_7)) from which our system is derived. Our improved system, called PATHE, includes (a) inexpensive (b) fully automated equipment (c) a system to implement selection for ***protein-protein*** binding, (d) facile methods for retuning the selective pressure to keep the affinity of the candidate ligand moving upwards through many rounds of mutagenesis, (e) new monitoring and feedback control methods, and (f) cassette-like vectors to facilitate ease of adaption to each new candidate ligand and target protein. An important distinguishing feature of PATHE is that unlike other selection systems the investigator does not need to presume which part of the candidate ligand to mutate. The automated PATHE system has the capacity to isolate a series of mutations that cumulatively improve the binding of a ligand to a target protein, at each phage generation (i.e. selecting from a population of *all possible double mutations* in the candidate ligand). ***Thus, the PATHE system will explore a much larger portion of the fitness landscape than other affinity selection systems.***

We intend to use this system to create a series of lead virulence factor inhibitor candidates that can inactivate pathogens regardless of their drug-resistance status. In particular, we will create proof-of-concept high affinity protein-based ligands that bind to two main pneumococcal virulence factors (CbpA & Pln) in the R21 phase and to fragments of both the A and B toxins produced by *Clostridium difficile* in the R33 phase. Biological systems have proven themselves to be extremely creative in acquiring genetic alterations that circumvent therapeutic agents. With resistance to ‘traditional’ antibiotics there is no opportunity to easily override resistance. However, in the event that ***mutation in the target protein*** ***leads to resistance***, new high affinity ligands can be readily created using the PATHE system using materials already in hand. Moreover the proposed studies will enable us to put a new high affinity selection system into the hands of a community of investigators which will expand development of protein-based therapeutics, biosensors, and diagnostics.

***It is also important to note that this technology will provide leverage to respond rapidly to new medical crises and emerging infections.*** These proof-of-concept studies will open new opportunities for other investigators. Once conditions have been negotiated for eliminating the regulatory barrier for probiotic therapy we expect that the protein-based virulence factor inhibitors created by this technology could be delivered *via* secretion by commensal bacteria to provide both prophylactic and therapeutic protection.

The creation of *pneumococcal* virulence factor ligands will justify progression to the R33 phase where we will document that these ligands are in fact inhibitors as ***hypothesized***. We will then use the PATHE system to create ligands to *C. difficile* toxins to further demonstrate the power of this technology.

***Data supporting this novel approach.***Mutations in virulence factors invariably have significant effects on the virulence of pathogens ([8-15](#_ENREF_8)). The consequences of mutations in each of the two virulence factors (CbpA & Pln), the object of our study, produced by *pnemucoccus* have been defined by the Tuomanen lab ([13](#_ENREF_13))*.* Those studies include mutations in both of the *pneumococcus* virulence factors we intend to inhibit, CbpA and Pln. These mutants showed a decreased ability to colonize the nasopharynx in the murine model system, indicating reduced virulence. Mutations in each also showed less-fulminate pneumonia than wild type. “Pln-deficient mutants exhibited a 5-log reduction in the number of pnemococci in the lungs” as well as reduced titers in the bloodstream and a diminished capacity to invade the CNS ([13](#_ENREF_13)). CbpA mutants were also deficient in the capacity to cross the blood-brain barrier. Consequently it is reasonable to ***hypothesize*** that protein-based inhibitors that bind to these two virulence factor inhibitors will also seriously reduce the virulence of the targeted pathogen regardless of the drug-resistance status of the bacteria.

Dr. Liu’s group at Harvard has utilized PACE to carry out a number of experiments involving evolution of T7 RNA polymerase ([7](#_ENREF_7), [16-18](#_ENREF_16)). Despite this exciting work, all of the publications concerning the use of PACE have come from the originating laboratory indicating that the PACE system is not easily exported to the research community. Nonetheless their studies establish the power of the basic technology from which we have derived the more user friendly PATHE technology.

***The PATHE system:*** Evolution of a new useful trait requires many generations of replication with concurrent mutagenesis and selection for the desired trait, in this case, binding to a target protein. The PATHE system (Fig. 1) has three basic components (a) a target plasmid (TP) encoding the target protein (the bait); (b) an evolving M13 phage (EP) encoding the candidate ligand (the prey) to be evolved into a high affinity ligand to the target protein; and (c) a three-hybrid system (Figs. 2) that links phage viability to the binding affinity of the candidate ligand to the target protein. Briefly, to use the PATHE system an EP that expresses a gene for a candidate ligand is constructed (Fig. 1). A Target Plasmid (Fig. 1) is constructed that contains a gIII gene encoding pIII (a protein required for phage viability) that is under the control of a cAMP-dependent promoter. cAMP is produced by a bacterial three-hybrid systemmodified from the *B. pertussis* adenylate cyclase two-hybrid system([19](#_ENREF_19))(Figs. 2 & 5) in which cAMP production in an adenylate cyclase minus bacterium is dependent on ligand binding to the target protein thus bringing the T25 and T18 subunits of adenylate cyclase together to enable cAMP production. The more effectively the ligand binds to the target protein (Figs. 2 & 5), the more cAMP is generated, leading to more pIII expression, and thus more infectious phage. Enrichment for tight binding occurs because the medium in the lagoon turns over faster than the cells replicate (to prevent bacterial mutations from effecting the selection) yet slower than the phage replicate.

**Figure 1. The PATHE System.** Evolution takes place in lagoons (4 per apparatus). Selection for binding takes place because the number of viable phage (***blue rods with pili***) produced per bacterial cell (***orange rectangle***) is proportionate to the affinity with which the candidate ligand binds to the target protein via a three-hybrid system that makes cAMP (Fig. 3) which drives production of the pIII protein necessary for phage viability. A laser based turbidity monitor controls a valve which regulates when mid-log Phase uninfected cells and media are inputted to the lagoon from a turbostat. The rate of medium input to the lagoons is faster than the cells replicate to prevent cell mutations from affecting selection but slower than phage replicate so phage which encode higher affinity ligands come to dominate the population. The *circles* represent DNAs in the bacterial cell; TP = ‘target’ plasmid encoding the target protein; MP = plasmid encoding a mutagenic polymerase; EP = ‘evolving’ phage encoding the candidate ligand protein.



**Figure 2.** **Three hybrid cAmp production selects exclusively for enhanced candidate ligand affinity.** A phage protein necessary for M13 viability is encoded by gene gIII which is driven by a cAMP dependent promoter. **The Two Hybrid System.** In this system when both the T18 and T25 ‘half’ modules of adenylate cyclase are brought together by binding between the target protein and the candidate ligand6, and cAMP is made. However, T18 and the candidate ligand are both encoded in the ‘evolving’ phage and hence are both subject to mutation (***green and red Xs***). The red mutations will accumulate until the T18 module binds to the T25 module of adenylate cyclase without any ligand protein binding to the target protein. **The Three-Hybrid System.** In this system only the candidate ligand and a leucine zipper tail are encoded in the ‘evolving’ phage and hence subject to mutation (***green Xs***). Loss of the leucine zipper by mutation gives no false positives.

***Why a Three-Hybrid System?*** It is essential to enforce selection only on the candidateligand**.** In a two-hybrid system (Fig.2), both the candidate ligand and the T18 module of adenylate cyclase would be in the phage that is undergoing mutagenesis and selection. This means that the T18 module would collect mutations and selection would find mutations that allow the T18 module to bind to the T25 module in the absence of binding between the candidate ligand and the target protein.

***Data supporting our capacity to implement PATHE***. (i) We have constructed a prototype apparatus to run PACE or PATHE based on the original publication ([7](#_ENREF_7)). (ii) That system maintains a sterile culture for at least two weeks. (iii) We have made the adenylate cyclase minus mutants necessary for the three hybrid system (Fig. 2) ([19](#_ENREF_19)). (iv) We have demonstrated evolution in our apparatus. After a 65 hour run, simply selecting for growth, every sequence unnecessary for phage growth was removed from the input phage (about 2 kb removed). (v) We have recapitulated the Harvard PACE study that converted a phage containing a T7 RNA polymerase that recognized a hybrid T7/T3 promoter into a polymerase that could recognize a pure T3 promoter after 99 hours of evolution (see Fig. 3). (vi) In addition, we have made the proteins necessary to measure binding affinities via isothermal titrating calorimetry. (vii) We have tested the various remote sensors that will be used in the inexpensive version (~$2,000) of the apparatus to do PATHE as described in ***Aim1A****.* These studies demonstrate that we have established the PACE technology in our laboratory.

****

**Figure 3. Evolution of a T7 RNA Polymerase to Recognize a T3 Promoter**. The blue and red lines illustrate the number of phage in the apparatus lagoons where evolution occurs. The blue line represents the number of phage that will utilize a hybrid promoter between T7 and T3. We start with a hybrid promoter to get phage that can evolve and then switch to the T3 promoter alone after 20 hours. The red line represents the number of phage that will utilize a pure T3 promoter.

**APPROACH: R21 PHASE**

Overview: We propose to implement and optimize a new, powerful affinity selection system that we will use to create high affinity protein-based ligands to pathogen virulence factors. We ***hypothesize*** that a significant fraction of these ligands will function as virulence factor inhibitors. This **hypothesis** will be tested in the R33 phase of the project. Several limitations for the use of the prototype PACE device ([7](#_ENREF_7)) for the generation of high affinity ligands must be overcome by implementing (a) an easy to employ apparatus to carry out PATHE experiments, (b) vectors to facilitate the DNA constructions to do PATHE, and (c) determining the conditions necessary to manipulate the selection stringency during a PATHE run to achieve high affinity.

**AIM 1:** **Produce the materials and data needed to implement the PATHE system**

**Aim 1A: Build a fully automated, cost-effective , easy to use PATHE apparatus**

Overview: This aim will be implemented by Peter Reintjes, the Innatrix instrument maker. The apparatus we are currently using ([7](#_ENREF_7)) is (a) bulky, (b) expensive (~$40,000), (c) error prone during assembly, (d) difficult to operate and (e) requires continual monitoring for good performance. Control of the input of uninfected cells, media, enzyme inhibitors and inducers is crucial for the operation of PATHE. These factors control the level of phage production, mutagenesis and selection during each cycle of phage replication. Thus, in this sub-Aim, Innatrix will develop an inexpensive (~$2,000), user-friendly, proprietary hardware and software system that will enable dozens of generations of evolution per day in autonomous mode over the course of several weeks. “Real time” analysis of the state of phage production can be performed directly via luminescence or by an auto- sampling device to collect samples of the population at desired intervals for high throughput evaluation.

Approach:The new hardware system will be based on replacing expensive pieces of precision hardware with very inexpensive hardware where precision is regained by micro-computer control and feedback from remote sensors. Liquid levels (interface detection *via* camera), stirring activity (visual sensing *via* camera), temperature (determined *via* infrared sensor), the presence of leaks (detected by conductance), viable phage levels in the lagoons (luminescence levels measured by camera from luciferase which is fused to M13 pIII) and cell density in the turbostat (*via* laser beam penetration) will all be monitored using sensors external to the various liquids. Sterilization will require only autoclaving the glass vessels. The chemically resistant tubing will be sterilized using 1 N NaOH hence the tubing will not need to be disassembled for sterilization and reassembled thus eliminating a major source of error. Performance of the apparatus, referred to as EvoStat, will be monitored remotely using Internet-of-Things enabled connections ([20](#_ENREF_20)).

Expected results, potential pitfalls and solutions: We anticipate no significant problems in constructing a prototype EvoStat. Within the first 6 months we expect to build the apparatus and hence to replace our current cumbersome equipment for the bulk of our experiments. All remote sensor systems and temperature control systems have been tested in isolation by Peter Reintjes, our instrument maker. We now need to assemble and test how the parts work together. If the least expensive parts fail in these tests we will replace them with parts that are less inexpensive but still less expensive than what we are currently using. As an example, the current valves we use cost $179 each, the new EvoStat uses ~$0.50 3D-printed valves. If those turn out to be inadequate we’ll use commercial $7.00 valves.

**Aim1B: Create a modular vector system to facilitate generation of high affinity ligands**

Overview: This aim will be implemented by Dr. Peter Charles and Martha Collier and a new technician (TBD). Vector construction is the bottleneck for any affinity selection system. For PATHE one needs a vector (which we call the Evolving Phage or EP) encoding the candidate ligand (the prey). PATHE also needs a vector (which we call the Target Plasmid or TP) encoding the target protein (the bait). We also need an expression vector to produce the proteins we need for affinity measurements which will be purchased from LifeTechnologies. We intend to construct vectors where the candidate ligand and target protein sequences can be changed easily and reliably.

Approach:We will incorporate Gateway cloning sequences ([21](#_ENREF_21)) to facilitate easy and reliable cloning ([22](#_ENREF_22)) of new candidate ligand and target sequences into the base vectors. This is a very simple construction process 

**Figure 4 Vectors to Facilitate Construction of Target Protein Vectors for PATHE.** The target protein sequence, constructed commercially with flanking attB sequences (25 bp), are inserted into a donor vector using a lambda recombinase BP enzyme. The resulting pENTR plasmid is then combined with our TP-base vector containing the T25 and T18 domains of adenylate cyclase and the gIII gene driven by a cAMP dependent promoter. In the TP-base vector the T18 domain has a leucine zipper (LZ) tail so that it will bind to the candidate ligand protein which will also be constructed with a complimentary leucine zipper tail. An LR recombination reaction will then replace the ccdB gene with the target protein gene in such a way that theT25 domain in the TP-base vector is fused to the target protein gene in the TP vector. In both the BP and LR reactions the desired sequence replaces the ccdB gene. ccdB is a lethal gene that targets DNA gyrase except in special strains. Any DNA molecule that has not had the ccdB gene removed by recombination will kill the recipient cell upon transformation and hence there is essentially no background to these reactions.

which has been proven by others to be robust and reliable (22) with only two enzymatic reactions and two transformations with essentially no background. An even simpler process than that in Fig. 4 is available to us for the incorporating the target sequence. The target sequence can be made commercially surrounded by attL sequences (100 bp each) and inserted into the TP-base vector directly with only one enzymatic LR reaction.

An Evolving Phage vector will be constructed in a similar fashion. The base vector will be an M13 phagemid missing the M13 gIII gene. After the LR reaction the plasmid origin of replication will be removed by the recombination event leaving an M13 defective phage with the candidate ligand gene fused to a leucine zipper. We will utilize Gateway vectors constructed by Dr. Ladant ([23](#_ENREF_23)) as the starting point of our design. These base vectors will be designed by us but synthesized commercially by Genewiz. The recombinase enzymes will be purchased from LifeTechnologies.

Expected results, potential pitfalls and solutions: We do not anticipate any significant problems with the construction of these vectors. The technology of utilizing Gateway cloning sequences is well established ([22](#_ENREF_22)). Since a commercial source will be used for construction this aim should proceed quickly.

**Aim1C: Develop Standard Operating Procedures for creating high affinity ligands to the target protein.**

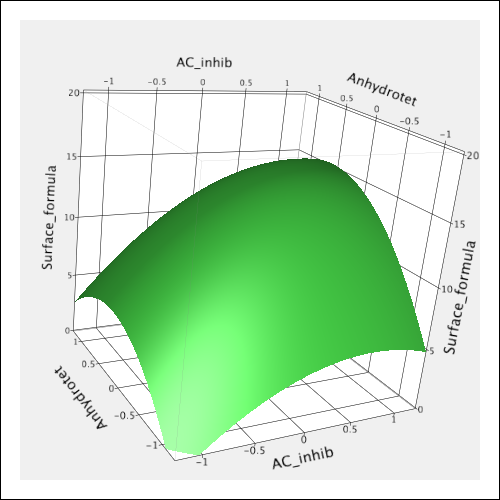
Overview: This aim will be designed and the data analyzed by Dr. Charles Carter. The data generation will be carried out by Dr. Peter Charles and Martha Collier and a new technician. The PATHE system will utilize a three-hybrid system to couple phage production to binding events between the candidate ligand and the target protein. Thus the more binding events that occur, the more phage are produced until production saturates the cell capacity (usually about 200 phage per cell). The affinity attained when the maximum number of phage are produced determines the upper limit on the affinity of the ligand to the target protein under those conditions. When the system saturates the first time it is unlikely that the ligand affinity will be sufficient for our purposes. At this point we need to reduce the number of phage produced per binding event to enable the production of higher affinity ligands. This sub-aim will map the pIII response surface ongoing within the PATHE apparatus to determine the temperature, adenylate cyclase inhibitor, and various gene inducer concentrations necessary to increase the stringency of the selection. At the saturation point we will want to reduce the phage output to ~10 per cell. We expect to increase the affinity from micromolar to nanomolar, or 1000 fold. If the phage yield is linear with affinity then we will need to make the transition from 200 phage per cell to 10 phage per cell ~50 times. We will write software to detect saturation and to inject the proper amounts of controlling reagents determined by this aim to reduce the output to 10 phage per cell.

Approach:Without knowing how to manipulate stringency, attaining high affinity would be hit-or-miss. We will therefore determine jointly how temperature, and concentrations of inducers (IPG), arabinose (Ara), and anhydrotetracycline (anhydrotet) as well as adenylate cyclase inhibitor (AC\_Inhib), (Fig. 5) influence phage production, just as Dr. Carter has previously optimized a multi-parameter RNA editing reaction ([24](#_ENREF_24)) using Response Surface methodology. Experimentally mapping a 5D response surface poses challenges because the goal is to accurately define the curvature of the targeted result (the “response”) with changes in each variable, and how that response changes from low to high levels of each of the other variables (two-way interactions). An appropriate design of experiments minimizes these challenges. An appropriate response surface experimental design distributes experimental tests both randomly and in a balanced way, such that the tests are located at points expected to have the greatest influence on the shape of the surface. Generally, this means several tests near the expected optimum, with the rest distributed as far away from the optimum as is convenient without losing the local quadratic approximation. This procedure avoids having to replicate simultaneous tests with each variable and to assess cross-correlations. A multivariate quadratic polynomial (Fig. 6) is an excellent local approximation to experimental behavior of many systems. Polynomial coefficients are evaluated by linear multivariate least squares fitting to the data. The gradient of the resulting analytical expression can be set to zero to solve for the optimized experimental conditions, and can then be used to estimate system behavior at different ligand-target affinities. Success of response surface methodology is due largely to the fact that because each experiment has been selected to maximize its influence on each of the coefficients in the context of all other data points, it does more work in the ensemble. A minimum of three experiments is required for each independent variable, especially if one is near the optimum and is matched by experiments far enough away from the optimum to detect the decreased value accurately. Further, the number of significant two-way interactions, hence the number of coefficients, generally increases more slowly with dimensionality than does the total number of possible two-way interactions. Finally, strong interactions can sometimes be transformed into more useful variables using principal components analysis. Computational simulations using two different strategies for experimental design, JMP ([25](#_ENREF_25)) and Gosset ([26](#_ENREF_26)) both show that rather high replication (3-4fold) is essential to defining optimal parameters with minimum integrated error. The success of response surface experiments is highly dependent on the range of responses given by the system.

A 5D quadratic polynomial (Fig. 6) requires at least 21 experiments to fit all coefficients. Initial tests will eliminate variables without effect on pIII production. We will use a design with 28 experiments prepared using Gosset ([26](#_ENREF_26)) and test a response surface with five experimental variables,. Each experiment will be replicated three times, to minimize the impact of experimental noise. Significant coefficients for the quadratic polynomial will be selected using stepwise multiple regression to test the inclusion or rejection of a given coefficient. Tests from simulations show that it will be possible to identify conditions whose response is within 2-5% of the theoretical optimum.



**Figure 5. Parameters Effecting the Number of Phage Produced per Prey/Bait Binding Event.** The affinity between the ligand (prey) and the target protein (bait) depends on when the cells saturate in terms of making phage. That prey/bait affinity at which saturation occurs can be controlled by the following five parameters: IPTG, temperature, adenylate cyclase inhibitor, arabinose, and anhydrotetracycline. The object is to control the number of phage produced per binding event. Phage production is limited by the amount of pIII produced. The amount of pIII produced depends on the amount of cAMP and the amount of the arabinose and anhydrotetracycline inducers in the cell which activate transcription of the ara/Tet cAMP dependent promoter driven gIII gene. The host cells are adenylate cyclase minus and hence produces no cAMP. This cAMP dependent ara/Tet promoter has a 1000 fold dynamic range depending on the amount of inducers present. The amount of cAMP in the cell depends on the amount of adenylate cyclase (AC) activity produced. The amount of AC activity per molecule of AC can be controlled by how much AC inhibitor is added to the medium and by the temperature at which the cells are growing. The number of AC molecules made will depend on how many candidate ligands bind to the target protein in a three-hybrid system (Fig. 3). The amount of target protein in the cell depends on a cAMP independent lacUV5 promoter with a 100 fold dynamic range in response to inducer concentration.

****

**Figure 6.** **Simulated 5-variable quadratic polynomial response surface.** The surface: [RS=0.5\*Temp1.5\*Temp2+1.3\*IPTG-2.0\*IPTG2+1.7\*AC\_Inh-2.0\*AC\_Inh2+1.9\*Ara-2.5\*Ara2+1\*Anhdrotet-3\*Anhodrotet2+0.5\*Temp\*IPTG+0.5\*Temp\*AC\_inh+1.2\*Temp\*Ara+0.4\*Temp\*Anhydrotet+0.8\*IPTG\*AC\_inh+0.2\*IPTG\*Ara+0.7\*IPTG\*Anhodrotet+0.3\*AC\_inh\*Ara+0.8\*Ac\_inh\*Anhodrotet+0.3\*Ara\*Anhydrotet]

isprojected onto its two most influential variables, the concentrations of AC inhibitor and inducer, Anhydrotetracycline. It was constructed to simulate recovery of coordinates of the optimal point from data for different experimental designs and can be recovered from 28 experimental datapoints. Location of the optimum at [Temp, IPTG, AC\_Inh, Ara, Anhydrotet] = [0.73,0.68,0.79.0.66,0.43], near the right-rear corner makes this surface more challenging to map experimentally than were the optimum centered near [0,0,0,0,0].

Expected results, potential pitfalls and solutions: This sub-aim is likely to take the most time during the R21 phase. For a given system, the integrated parameter error has three important determinants: the average error of the measurement ([27](#_ENREF_27)), the experimental design strategy, and their two-way interaction. Simulations with perfect and randomly perturbed datasets to test exactly the same multivariate quadratic model show that a minimum prediction variance design on a sphere is significantly more robust to experimental noise than the central composite design afforded by JMP for the same number of experiments. Furthermore, such designs also work better as the experimental noise increases, the preferential effect is greater, the greater the error in the measurements. Data files and algorithms associated with this model will enhance designs and aid in troubleshooting unforeseen problems arising during the optimization of runtime values for the five independent variables that control pIII expression. Use of luminescence as a metric should accelerate throughput, compared with counting plaque forming units, enabling us to do response surface experiments at different ligand-target affinities. However, a potential pitfall is that the variance in the phage measurements could be too large given our inexpensive luminescence sensor. If we are unable to reduce the variance sufficiently (to a few percent) with our apparatus we’ll buy a commercial apparatus and modify it for high throughput measurements as Dr. Edgell has done before ([28](#_ENREF_28)) with an automated titrating fluorimeter.

\

**AIM 2: Construct high affinity protein-based ligands to two of the major virulence factors (CbpA & Pln) produced by *Streptococcus pneumonia*e**

**Overview:** *Streptococcus pneumoniae* produces six virulence factors ([13](#_ENREF_13)); the main two factors are choline binding protein A (CbpA) responsible for binding to cells ([12](#_ENREF_12)) and pneumolysin (Pln) responsible for pore formation in the target cells ([13](#_ENREF_13)). The function of each of these virulence factors has been determined by mutagenesis followed by quantization of bacterial titers in relevant organs and real-time bioluminescent imaging ([13](#_ENREF_13)). We will use the PATHE system developed in ***Aim1*** to generate 5-10 high affinity protein-based ligands that bind to each of these two *pneumococcal* virulence factors. In this aim we will employ two different generic protein binding scaffolds for ligand evolution, single chain antibody fragments (scFv) and fibronectin monobodies to increase the probability of finding effective inhibitors in the R33 phase within this set of ligands.

**Aim2A: Create single chain antibody fragment (scFv) high affinity protein-based ligands to the two major *pneumococcal* virulence factors (CbpA & Pln).**

Overview: This aim will be implemented by Drs. Edgell & Charles, and by Martha Collier and a new technician. The scFv protein family represents a generic scaffold with the capacity to be mutagenized to bind to a very large range of epitopes ([29](#_ENREF_29)). The PATHE system evolves phage based on selection for binding to the target protein. To have phage that can evolve we need to start with a candidate ligand that binds at least modestly well to the target protein. A commercial company, AxioMx (<http://axiomxinc.com>) will produce scFv proteins binding to the target proteins that we provide *via* traditional phage display

Approach**:** The PATHE system requires target proteins that are soluble in the cytoplasm of *Escherichia coli*. These two *pneumococcal* virulence factors are both soluble in the *E. coli* cytoplasm ([12](#_ENREF_12), [30](#_ENREF_30)). Given a starter candidate ligand gene from AxioMx we will lift that sequence with primers that will surround the ligand sequence with attB sequences (21 bp). That candidate ligand gene will then be inserted into the Gateway Evolving Phage base vector produced in ***Aim1B*** using the Gateway BP reaction to move the candidate ligand to a pDONR plasmid followed by a Gateway LR reaction to move the candidate ligand gene from the pDONOR plasmid into the EP-base vector (very similar to Fig. 3). The target virulence factor genes surrounded by attL sequences (100 bp) will be synthesized commercially by Genewiz and inserted into the Gateway Target Plasmid created in ***Aim1B*** using a Gateway LR reaction (last reaction in Fig. 3). These two constructs will then be employed in a PATHE run to produce high affinity ligands to the target virulence factor.

A PATHE run will consist of growing the EP phage containing the candidate ligand of interest in a lagoon with bacteria at mid-log phase containing a plasmid encoding the target protein of interest (TP). The lagoon medium is exchanged with fresh medium and mid-log cells at a rate faster than the cells divide (to prevent bacterial mutations from impacting the selection) but slower than the phage replicate. Phage yield will be monitored via the luminescence provided by a luciferase enzyme fused to the M13 pIII protein. When the computer determines that the phage yield has plateaued this indicates that the cells are producing as many phage as possible given the growth conditions. Initially the binding affinity will not be as high as desired due to phage production saturation. At this point the stringency determining the number of phage produced per binding event will be increased by the computer using the data generated in ***Aim1C***. This will be repeated up to 50 times until the phage production no longer increases when the stringency is increased. This would indicate that we have reached the highest binding affinity attainable with these components. At this point the ligand DNA will be sequenced. The target ligand sequence will then be isolated by PCR and inserted into a Gateway expression vector. Expressed protein will then be isolated by His-tag chromatography and its affinity for the target protein measured by isothermal titrating calorimetry in the UNC Macromolecular Interactions Facility (<http://www.med.unc.edu/csb/macinfac>).

Expected results, potential pitfalls and solutions: A successful run will be indicated by a high yield of phage under a stringent phage per binding event condition defined in ***Aim1C***. and a population containing 5-10 different ligands A potential limitation will be if the virulence factor ligands do not attain affinities in the nanomolar range. The solution for this problem will be to employ a different generic scaffold (e.g. fibronectin monobodies) … which will be pursued in ***Aim2B*** *&* ***Aim2C*** outlined below. If nether works intermedilysin has 50% identity to pneumolysin and binds to the cell receptor CD59 (76 aa) ([31](#_ENREF_31)). If we have trouble with pneumolysin we could use CD59 as a starter scaffold. We expect that it will require approximately three months to produce and characterize the two ligands to virulence factors.

**Aim2B: Create fibronectin monobody high affinity protein-based starter ligands to two major *pneumococcal* virulence factors (CbpA & Pln).**

Overview: Dr. Brian Kuhlman, who has experience at creating monobody ligands ([32](#_ENREF_32)) will be responsible for this aim. To increase the likelihood that inhibitors will be found within the set of ligands that bind to *pneumococcal* virulence factors we will create high affinity ligands using two different generic protein binding scaffolds. The fibronectin monobody represents a generic scaffold with the capacity to be mutagenized to bind to a very large range of epitopes ([33](#_ENREF_33)). This sub-aim is designed to produce starter ligands using the fibronectin monobody generic binding scaffold.

Approach: To generate initial binders against the two virulence factors we will use standard phage display ([34](#_ENREF_34)) with FNIII monobodies as our binding scaffold. Monobodies have been successfully used to target a wide range of proteins including kinases and membrane proteins ([35](#_ENREF_35)) ([32](#_ENREF_32)). We will use combinatorial libraries that vary residues in the FG and BC loops of the monobody (Fig. 7). The loops resemble the CDR loops of antibodies in that they are tolerant to a wide range of mutations. Seven residues will be varied in the FG loop and five residues in the BC loop. To keep the size of the library on a scale that can be sampled with phage display (~109 variants), several positions in both loops will be restricted to only sampling tyrosine and serine by using the degenerate codon TMY. Previous studies have shown that tyrosine and serine are highly enriched at antibody/antigen interfaces and that libraries only containing serine and tyrosine are often sufficient to generate moderate affinity binders ([35](#_ENREF_35)). The monobodies will be expressed on the surface of phage as a fusion with the pIII.

**Figure 7.** **Structure of FNIII monobody** (PDB: 1FNA) with the BC loop colored in blue and the FG loop in red. **

To enable biopanning of the phage library, the virulence factors will be biotinylated (via lysine side chains) and captured on streptavidin magnetic beads. The phage will then be incubated with the beads and non-binders will be removed with multiple washes. The bound phage will be eluted with a low pH solution and then used to infect cells in order to amplify the population. Three to four rounds of selection will be used to identify binders. Following the identification of unique clones via DNA sequencing, the FNIII variants will be cloned into a pGEX-4T-1 vector for expression as GST fusion proteins in bacteria. The GST-FNIII proteins will be purified with affinity chromatography, and following protease cleavage to remove the GST will be further purified with size exclusion chromatography. Isothermal titration calorimetry (ITC) will be used to verify binding between the FNIII domains and their cognate virulence factors. ITC is a robust strategy for detecting binding events with dissociation constants between 100 nM and 10 μM, a common affinity range for first generation monobodies generated with phage display ([36](#_ENREF_36)).

Expected results, potential pitfalls and solutions: A successful run will be indicated by a isolating a monobody binder in the 100 nM to 10 μM range. We do not anticipate difficulties with this sub-aim since the co-PI who will be doing these experiments has already been successful at generating modest affinity ligands to two other target proteins ([32](#_ENREF_32)) using fibronectin monobodies.

**Aim2C: Affinity maturate the fibronectin monobody starter ligands to the two *pneumococcal* virulence factors (CbpA & Pln).**

Overview: This aim will be implemented by Drs. Edgell & Charles, and by Martha Collier and a new technician. Given modest binding ligands to the target proteins we expect to be able to use PATHE to evolve variants with much higher affinity, that is, nanomolar or sub-nanomolar.

Approach: The approach here is identical to the approach in ***Aim2A***.

Expected results, potential pitfalls and solutions: A successful run will be indicated by a high yield of phage under the most stringent phage per binding event condition defined in ***Aim1C***. If the technology is working this should proceed easily. Even if the PATHE runs need to be carried out for two weeks the entire set can be generated and characterized in 3 months. A potential pitfall here is that the fibronectin monobody ligands might not attain high affinity (nanomolar) binding to the virulence factors. Because we are using different generic scaffolds (e.g. scFv) … which is being pursued in ***Aim2A*** we do not anticipate any difficulties in generating at least 10-20 candidate ligands for testing for inhibition in the R33 phase.

**APPROACH: R33 PHASE**

**AIM 3:** **Determine if the high affinity protein-based ligands binding to *pneumococcal* virulence factors (CbpA & Pln) function as inhibitors in mice as hypothesized.**

**Overview:** If the specific milestones of the R21 phase are met we expect to generate 10-20 high affinity protein-based ligands to two *pneumococcal* virulence factors. The central goal of the R33 phase is to determine if this set of high affinity ligands contains biologically functional virulence factor inhibitors. Using two different generic binding scaffolds in the R21 phase to create the virulence factor ligands increases the likelihood that we will find inhibitors within the set of protein-based high affinity ligands.

**Aim3A: Inhibition of interaction of CbpA with Laminin Receptor (LR) *in vitro*.**

Overview: This aim will be implemented by Dr. Marcia Hobbs in close collaboration with Dr. Elaine I. Tuomanen. We will test the virulence factor ligands for inhibitory activity in ELISA and cell culture based systems *in vitro*. These assays reflect activity of the virulence factors in natural disease ([13](#_ENREF_13)).

Approach: CbpA is known to interact with LR on the surfaces of endothelial cells during the course of invasive pneumococcal infection. ([14](#_ENREF_14)) Two *in vitro* assays will be used to test inhibitors: inhibition of direct binding of pneumococci (wild type and *cbpA-*) to LR in an ELISA format and inhibition of bacterial binding to endothelial cells in culture. In the ELISA assay, purified rLR or BSA (2-50 μg/ml) are used to coat amino-reactive 96‑well microtiter plates. Digoxigenin (Roche) labelled bacteria (wild type TIGR4 or a *cbpA-* mutant) are added to the wells for 2-4 h at room temperature. Plates are washed and bound bacteria are detected with polyclonal anti-digoxigenin Fab fragment (POD)‑conjugated antibody (1:5000; Roche). Inhibition assays will be performed with bacteria pre-incubated with 20 μg/ml of candidate inhibitor produced by his-tag chromatography in aim 2A and aim 2C or positive control rLR. Effective inhibitors will then be characterized with a dose response curve.

Adherence of bacteria will then be modeled using rat brain capillary endothelial cell line rBCEC6 activated with TNFα (10ng/ml).Cells are incubated with 1 × 107 cfu/ml bacteria, washed and adhesion is quantitated by trypsinizing the cells and plating for CFU.In inhibition experiments, control LR peptides include LR amino acids 263-282 (blocks adherence) and a scrambled version of 263-282 (fails to block).To test blocking of bacterial adhesion to rBCEC6 endothelia, 107 cfu bacteria will be pre-incubated with candidate protein ligand or control peptide for 1 h at room temperature. Bacteria will be washed and then run through the adherence assay. We will test a range of ligand concentrations to generate dose response curves for candidate inhibitors and the positive control LR peptide and determine the molar concentrations that inhibit 50% of bacterial adherence (effective dose, ED50). Candidate inhibitors with an ED50 equal to or less than the positive control will move on to mouse infection studies, as described in Aim 3C.

**Aim3B: Inhibition of pneumolysin cytotoxicity *in vitro*.**

Overview: This aim will be implemented by Dr. Marcia Hobbs in close collaboration with Dr. Elaine I. Tuomanen. Pneumolysin is a cholesterol dependent cytolysin that oligomerizes to form large pores in eukaryotic cell membranes thereby lysing cells. This results in death of virtually all cell types and is the basis of its toxicity during the course of infection. Toxicity can be quantified by demonstrating lysis of mammalian cells *in vitro*, including red blood cells or epithelial cells ([37](#_ENREF_37)).

Approach: The threshold concentration for hemolytic activity is determined by mixing serial dilutions of 10 g/ml pneumolysin with rabbit red blood cells and examining the mixture for hemolysis spectrophotometrically (OD540). Candidate proteins will be added to each well of red blood cells prior to addition of pneumolysin to test inhibition of activity. Positive controls for toxin activity are wild type pneumolysin without the candidate ligand or 1% Triton X-100. We will test a range of ligand concentrations to generate dose response curves for candidate inhibitors and determine the molar concentrations that inhibit 50% of bacterial lysis (effective dose, ED50). Inhibitors with the lowest ED50 will be tested with lung epithelial cells.

Because *pneumococci* infect the lung, cytotoxic activity can also be determined by overlaying pneumolysin on A549 lung epithelial cells and examining the viability of the cells using a MTT cell proliferation assay. A positive control for toxin activity, wild type pneumolysin, is run simultaneously. In 96 well plates, A549 cells are seeded at 3.5 x 105 cells/ml at 200uL /well in serum free DMEM/F12. Pneumolysin is added one day later at 10ug/mL in dilution buffer in well 1 (200L volume alone or with 1:2 serial dilutions of candidate inhibitors at concentrations surrounding the ED50 determined with rabbit RBCs. After 24 hours at 37°C in 5% CO2 A549 cell viability will be measured using the Vybrant MTT Cell Proliferation kit (Molecular Probes). The ED50 for A549 cells will be determined, and inhibitors with the lowest effective doses will move on to *in vivo* studies.

**Aim3C: Pneumococcal challenge of mice:**

Overview: This aim will be implemented by Dr. Marcia Hobbs in close collaboration with Dr. Elaine I. Tuomanen. Up to 6 inhibitors showing the most promise in either the CbpA or pneumolysin *in vitro* assays will advance to testing in mice for the ability to arrest disease.

Approach: For the pneumonia infection model, mice are challenged with 1x107 cfu of wild type *pneumococcus*, TIGR4X intratracheally ([13](#_ENREF_13)). Candidate protein inhibitors will be incubated with bacteria in the challenge inoculum. The TIGR4X strain is marked with a bioluminescent marker that is detected by Xenogen imaging. Thus the course of infection can be followed in real time without sacrifice of the mouse. Generally, the infection spreads from the lungs to the blood stream in 48 hours and the mouse is euthanized when moribund at 72 hours. An effective inhibitor would delay or prevent spread to the bloodstream and lethality. Inhibitors will be tested in 5 mice/group with untreated mice as controls. To confirm imaging data, blood cultures will be taken twice daily (mice can be sampled multiple times per day without sacrificing the animals). The course of disease in the presence of inhibitor will be compared to bacterial mutants that are deleted for either CbpA or pneumolysin genes that represent complete loss of function of these virulence determinants. Initial *in vivo* experiments will be conducted using the ED50 as determined in the *in vitro* studies with endothelial cells for CbpA and A549 cells for pneumolysin as described above. Escalating and/or multiple doses will be tested in the event bacterial burden or disease course in the presence of inhibitors is not statistically significantly impacted by the inhibitors.

Expected results, potential pitfalls and solutions**:** The *in vitro* and *in vivo* assays are standard in the Tuomanen lab and have been used by others successfully. Thus the techniques should not be a problem. For *in vitro* assays, it is possible that candidate protein-based ligands will have unintended biological activities such as killing endothelial cells or that they will be insoluble in culture media. This will result in exclusion of the ligand from further consideration. *In vivo* assays are anticipated to be a major challenge test for the candidate ligands. The pharmacodynamics of the protein-based ligands will not be known and it is possible that these molecules will be cleared so quickly that a sustained effect on disease over days will not be maintained. If so, then multiple intratracheal doses may be required. Alternatively, the model will be altered to focus on blood stream events which are much more rapid. Bacteria will be given IV and the course of infection including the development of meningitis will be followed. Multiple IV doses of candidate ligand can be tested. The proposed *in vivo* studies will test whether newly developed inhibitors can prevent or reduce the severity of pneumococcal infections in mice. Successful demonstration of such capacity will galvanize efforts to optimize inhibitors to treat existing infections in future studies.

**AIM 4. Document that the PATHE system can produce high affinity ligands relatively easily by creating high affinity protein-based ligands directed at two toxins produced by *Clostridium difficile***

***.***

Overview**:** At this point the PATHE technology should be fully implemented. Hence it should be relatively straight-forward to create high affinity protein-based ligands to any protein soluble in the cytoplasm of *Escherichia coli*. *Clostridium difficile* is rated by the CDC as an Urgent Health Threat ([2](#_ENREF_2)).

***CDC Report Information on C. difficile***. ([2](#_ENREF_2)). “Clostridium difficile (C. difficile) causes life-threatening diarrhea. These infections mostly occur in people who have had both recent medical care and antibiotics. Often, C. difficile infections occur in hospitalized or recently hospitalized patients.

RESISTANCE OF CONCERN

■ Although resistance to the antibiotics used to treat C. difficile infections s not yet a problem, the bacteria spreads rapidly because it is naturally resistant to many drugs used to treat other infections.

■ In 2000, a stronger strain of the bacteria emerged. This strain is resistant to fluoroquinolone antibiotics, which are commonly used to treat other infections.

■ This strain has spread throughout North America and Europe, infecting ankilling more people wherever it spreads.

PUBLIC HEALTH THREAT

■ 250,000 infections per year requiring hospitalization or affecting already hospitalized patients.

■ 14,000 deaths per year.

■ At least $1 billion in excess medical costs per year.

■ Deaths related to C. difficile increased 400% between 2000 and 2007, in part because of a stronger bacteria strain that emerged.

■ Almost half of infections occur in people younger than 65, but more than 90% of deaths occur in people 65 and older.

■ About half of C. difficile infections first show symptoms in hospitalized or recently hospitalized patients, and half first show symptoms in nursing home patients or in people recently cared for in doctors’ offices and clinics.” ([2](#_ENREF_2)).

Overview: ***This aim is identical to the R21 phase Aim2*** except that the target proteins here are the cell receptor binding domains of the *C. difficile* toxins A and B (TcdA and TcdB). These modules are those to which inactivating antibodies bind. They are small (142aa and 382aa) and soluble in the cytoplasm of *Escherichia coli.* ([12](#_ENREF_12), [30](#_ENREF_30)) The genes for these target proteins will be made commercially by Genewiz. This aim will be implemented by Drs. Edgell, Charles & Kuhlman and Martha Collier and a new technician.

The Approach & Expected results, potential pitfalls and solutions are identical to the three sub-aims in the R21 phase ***Aim 2***.

**Timeline and Interactions Between Participating Sites**

**Timeline:**

**R21 Phase:** Year 1: Aim 1A and Aim 1B; Year 2: Aim 1C and Aim 2

**R33 Phase:** Year 3: Aim 3 (initiation); Year 4: Aim 3 (completion); Year 5: Aim 4

**Interactions between participating sites:**

**R21 Phase:** Weekly project meetings between the co-principle investigators involved with aims 1 & 2 will be held at the University of North Carolina site. Semi-annual meetings of all of the Innatrix Scientific Advisory Board members will be held at UNC.

**R33 Phase:** Same as the R21 phase except we will add monthly Skype sessions with the consulting investigator at St. Jude involved with aim 3.

**MILESTONES FOR THE R21 and R33 PHASES**

**R21 milestones.** The proposed R21 milestones for determining the success of the R21 phase are:

Milestone #1. A cost-effective (20 fold reduced from the original description in ([7](#_ENREF_7))) functional apparatus for doing PATHE.

Milestone #2: Functional Gateway vectors for doing PATHE.

Milestone #3: A set of equations defining the dependence of phage yield per binding event as a function of the run parameters (temperature, inducer levels, adenylate cyclase inhibitor levels).

Milestone #4: 5-10 high affinity protein-based ligands against at least one of the two targeted *Streptococcus pneumoniae* virulence factors.

**R33 milestones:** The proposed R33 milestones for determining the success of the R33 phase are:

Milestone #5: Documentation that some of the high affinity ligands produced in the R21 phase that bind to *pneumococcal* virulence factors act as inhibitors in mice and the definition of their effects.

Milestone #6: 5-10 high affinity protein-based ligands that bind to the receptor binding domains (TcdA & TcdB) of the toxins produced by *C. difficile*.

**Future Directions:** If the R33 phase is successful the high affinity protein-based inhibitors will be developed further via collaboration in which we will seek out a company interested in furthering studies concerning clinical studies concerning efficacy, safety and toxicity, and ultimately distribution of the inhibitors. We will also seek funding to test the inhibitors in macaques. We will also seek to collaborate with an instrument company to sell our EvoStat and vectors for doing affinity selection.

**BIBLIOGRAPHY**

1. WHO (2014) Antimicrobial Reistance Global Report on Surveillance. (<http://www.who.int/drugresistance/documents/surveillancereport/en/)>.

2. CDC (2013) Antibiotic Resistant Threats. (Center for disease Control, <http://www.cdc.gov/drugresistance/threat-report-2013/)>.

3. Wilson DR & Finlay BB (1998) Phage display: applications, innovations, and issues in phage and host biology. *Canadian journal of microbiology* 44(4):313-329.

4. Husimi Y (1989) Selection and evolution of bacteriophages in cellstat. *Advances in biophysics* 25:1-43.

5. Husimi Y, Nishigaki K, Kinoshita Y, & Tanaka T (1982) Cellstat-a continuous culture system of a bacteriophage for the study of the mutation rate and the selection process of the DNA level. *The Review of scientific instruments* 53(4):517-522.

6. Aita T, Morinaga S, & Husimi Y (2005) Thermodynamical interpretation of evolutionary dynamics on a fitness landscape in an evolution reactor, II. *Bulletin of mathematical biology* 67(3):613-635.

7. Esvelt KM, Carlson JC, & Liu DR (2011) A system for the continuous directed evolution of biomolecules. *Nature* 472(7344):499-503.

8. Allen KP, Randolph MM, & Fleckenstein JM (2006) Importance of heat-labile enterotoxin in colonization of the adult mouse small intestine by human enterotoxigenic Escherichia coli strains. *Infection and immunity* 74(2):869-875.

9. Erume J*, et al.* (2008) Comparison of the contributions of heat-labile enterotoxin and heat-stable enterotoxin b to the virulence of enterotoxigenic Escherichia coli in F4ac receptor-positive young pigs. *Infection and immunity* 76(7):3141-3149.

10. Faherty CS*, et al.* (2012) Chromosomal and plasmid-encoded factors of Shigella flexneri induce secretogenic activity ex vivo. *PloS one* 7(11):e49980.

11. Harrington SM*, et al.* (2009) The Pic protease of enteroaggregative Escherichia coli promotes intestinal colonization and growth in the presence of mucin. *Infection and immunity* 77(6):2465-2473.

12. Luo R*, et al.* (2005) Solution structure of choline binding protein A, the major adhesin of Streptococcus pneumoniae. *The EMBO journal* 24(1):34-43.

13. Orihuela CJ, Gao G, Francis KP, Yu J, & Tuomanen EI (2004) Tissue-specific contributions of pneumococcal virulence factors to pathogenesis. *The Journal of infectious diseases* 190(9):1661-1669.

14. Orihuela CJ*, et al.* (2009) Laminin receptor initiates bacterial contact with the blood brain barrier in experimental meningitis models. *The Journal of clinical investigation* 119(6):1638-1646.

15. Zhang J*, et al.* (2013) Deletion of pic results in decreased virulence for a clinical isolate of Shigella flexneri 2a from China. *BMC microbiology* 13:31.

16. Badran AH & Liu DR (2015) In vivo continuous directed evolution. *Current opinion in chemical biology* 24c:1-10.

17. Dickinson BC, Leconte AM, Allen B, Esvelt KM, & Liu DR (2013) Experimental interrogation of the path dependence and stochasticity of protein evolution using phage-assisted continuous evolution. *Proceedings of the National Academy of Sciences of the United States of America* 110(22):9007-9012.

18. Dickinson BC, Packer MS, Badran AH, & Liu DR (2014) A system for the continuous directed evolution of proteases rapidly reveals drug-resistance mutations. *Nature communications* 5:5352.

19. Karimova G, Pidoux J, Ullmann A, & Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proceedings of the National Academy of Sciences of the United States of America* 95(10):5752-5756.

20. Lucas Martinez N, Martinez JF, & Hernandez Diaz V (2014) Virtualization of event sources in wireless sensor networks for the internet of things. *Sensors (Basel, Switzerland)* 14(12):22737-22753.

21. LifeTechnologies (GatewayCloning Technology. [*http://www.lifetechnologies.com/us/en/home/life-science/cloning/gateway-cloning.html*](http://www.lifetechnologies.com/us/en/home/life-science/cloning/gateway-cloning.html)*.*

22. Marsischky G & LaBaer J (2004) Many paths to many clones: a comparative look at high-throughput cloning methods. *Genome research* 14(10b):2020-2028.

23. Ouellette SP, Gauliard E, Antosova Z, & Ladant D (2014) A Gateway((R)) -compatible bacterial adenylate cyclase-based two-hybrid system. *Environmental microbiology reports* 6(3):259-267.

24. Chester A, Weinreb, V., Carter, C.W. Jr., and Navaratnam, N. (2004) Optimization of apolipoprotein B mRNA editing by APOBEC1 apoenzyme and the role of its auxiliary factor. *ACF.RNA* 10:1399-1411.

25. SAS (2004) JMP: The Statistical Discovery Software (SAS Institute, Cary, NC), V.6.

26. Hardin RH & Sloane NJA (1993) A new approach to the construction of optimal designs. *Journal of Statistical Planning and Inference* 37:339-369.

27. Edgell MH, Sims DA, Pielak GJ, & Yi F (2003) High Precision, High Throughput Stability Determinations Facilitated by Robotics and a Semiautomated Tiotrating Fluorimeter. *Biochemistry* 42(24):7587-7593.

28. Yi F, Sims, A.,Pielak,G.J., and Edgell, M.H. (2005) High-throughput, precise protein stability measurements facilitated by robotics. *J. Assn. Laboratory Automation* 10:98-101.

29. Nelson AL (2010) Antibody fragments: hope and hype. *mAbs* 2(1):77-83.

30. Mitchell TJ, Walker JA, Saunders FK, Andrew PW, & Boulnois GJ (1989) Expression of the pneumolysin gene in Escherichia coli: rapid purification and biological properties. *Biochimica et biophysica acta* 1007(1):67-72.

31. Johnson S, Brooks, N.J., Smith, R.A., Lea, S.M., and Bubech, D. (2013) Structural basis for recognition of the pore-forming toxin intermedilysin by human complement receptor CD59. *Cell Rep.* 3(5):1369-1377.

32. Gulyani A*, et al.* (2011) A biosensor generated via high-throughput screening quantifies cell edge Src dynamics. *Nature chemical biology* 7(7):437-444.

33. Bloom L & Calabro V (2009) FN3: a new protein scaffold reaches the clinic. *Drug discovery today* 14(19-20):949-955.

34. Sidhu SS, Lowman, H.B., Cunningham, B.C.,& Wells, J.A. (2000) Phage display of selection of novel binding peptides. *Methods in Enzymology* 328.

35. Koide A, Gilbreth, R.N., Esaki, K., Tereshko, V., and Koide, S. (2007) High-affinity single-domain binding proteins with a binary-code interface. *PNAS* 104(16):6632-6637.

36. Koide AK, S. (2007) *Monobodies* (Humana Press) pp 95-109.

37. Mann B*, et al.* (2014) Broadly protective protein-based pneumococcal vaccine composed of pneumolysin toxoid-CbpA peptide recombinant fusion protein. *The Journal of infectious diseases* 209(7):1116-1125.