The Applications of High Throughput Screening and its Value as a Drug Discovery Tool in the Pharmaceutical Industry and Academia

Presented to the S. Daniel Abraham Honors Program in Partial Fulfillment of the Requirements for Completion of the Program

Stern College for Women

Yeshiva University

April 27, 2021

Alexandra Huberfeld

Mentor: Dr. Alyssa Schuck

Table of Contents:

Abstract

High throughput screening is a method of drug development that has become increasingly popular in the past three decades. Since its popularization in the 1990s, it has been criticized by proponents of typical drug development strategies as costly and ineffective. This paper outlines the technique of high throughput screening, its history, and present uses. Using the design of a screen to identify small molecule activators for a sodium-citrate transporter in order to treat SLC13A5 Deficiency, the process of high throughput screening is outlined. The paper argues that while investing in the process is valuable for large pharmaceutical companies, it is not justifiable for academic researchers.

Introduction

Classical Pharmacology

Historically, drugs have been developed in several different ways. The first is through the discovery of natural products and adaptation towards drug use. An effect of a drug is first observed in a biological system, and then an effort is made to understand why the drug has that effect. This is known as classical pharmacology. The discovery of the antibiotic penicillin by Alexander Fleming in 1928 is an example of a drug developed through a product created naturally. Often, as with penicillins, these types of discoveries are serendipitous, first relying on luck to uncover a natural product that has the desired consequence. Once the organic compound is identified, then it can be fine-tuned and made into a drug for wide production and made safe to be consumed. Through nearly a century of research and development, penicillins are now known as an entire class of drugs, subdivided into five groups, each with a different range of targets and effectiveness. (ACS, n.d.) For example, Penicillin G was the original antibiotic discovered by Fleming, but another compound, Penicillin V was isolated from the same strain of fungus and was found to be more acid-resistant, thus more compatible for oral consumption. Drugs developed from natural compounds are the results of lengthy trial and error and often take years, even decades, to reach the market. (Medline, n.d.)

Another example of a drug class originally derived from nature is that of opiates. Originally isolated from poppy flowers, opium was used for centuries for pain relief and recreational use. However, as our understanding of opium increased, different chemicals were

isolated and adapted. Morphine, codeine, and heroine are all different versions of opium that were developed before the 18th century. Then Methadone, OxyContin, and Percocet were synthesized. Each iteration attempted to prevent the disastrous effect of addiction and increase the potency. (Foundation for a Drug Free World, n.d.) The history of opioids reflects a similar pattern of discovery to penicillin– a fortuitous natural discovery followed by a lengthy development period.

Reverse Pharmacology

The classical pharmaceutical approach described previously is not the only path to creating a successful drug. Another approach is called reverse pharmacology, in which the desired target is known, and chemicals are either synthesized or discovered to bind to the target. This target is usually some sort of protein involved in the desired pathway; a receptor or enzyme could be the object of this search. If the target is well known, researchers can engage in "rational design" in which a small molecule is created to bind to a target. By knowing the three-dimensional shape of a protein target, small molecules can be designed to specifically interact with it. This interaction can cause it to be activated, inactivated, or may have no effect whatsoever. (Mavromoustakos T. *et al,* 2011)

However, if a target is not well-mapped, less exact methods need to be utilized. This involves exposing a target to a wide array of potential ligands, with the objective that some will successfully bind. This array of molecules includes synthetic molecules, but also features natural molecules, like ones used in classical pharmacology. Molecules that have been found to

successfully bind to the target are subject to further screening to determine if any of the ligands have the desired physiological effect, often using cellular and then animal models. If successful, the drug can be evaluated in a clinical trial, and then expanded for public use. This process involves much trial and error, often taking years from the time a hit is identified until it is available for the market. (Arulsamy *et al.,* 2016)

High Throughput Screening

Reverse pharmacology has become more popular in recent decades, as the time-consuming process of classical pharmacology can fail to yield viable drugs even after decades of trial. In particular, screening technologies have become standard industry tools because it is more straightforward pathway to drug development, and the automation allows for hundreds of thousands of compounds to be screened in a short amount of time. The process seemingly promises results as long as enough chemicals are tested. This has led many pharmaceutical companies and universities to invest in screening as a drug discovery method and to amass enormous chemical libraries to be used in screens. Initially, the size of chemical libraries was between 10,000 to 100,000 compounds, however, in the early 2000s, some libraries reached 500,000 compounds (Maccaron *et al.*, 2011).

High throughput screening is a technique used in reverse pharmacology drug development in which automated equipment is used "to rapidly test thousands to millions of samples for biological activity at the model organism, cellular pathway, or molecular level." (Attene-Ramos *et al.*, 2014). What makes a regular screen into a "high throughput" one is hard to

define, but “is often defined by the rate, quantity, or automation level of the testing.” (Janzen, 2014) High throughput screening is a more advanced method of screening, which relies on the use of automatic equipment, instead of being performed by an individual or team of researchers.

History of High Throughput Screening

High throughput screening was developed by large pharmaceutical companies like Pfizer in the mid-late 1980s. In the course of developing a screening system of antibiotics against actinomycetes, bacteria in soil, the researchers at Pfizer wanted to increase the rate at which antibiotics could be tested for effectiveness. The lab was able to screen 10,000 samples per week in 1990, up from 800 per week in 1984. This was accomplished by the implementation of several different new technologies, including multi-well plates, up to 96-well plates, and multi-well pipettes. This increased the yield and rate of the screen, and this process was quickly expanded to other targets (Pareira and Williams, 2007). Other companies assembled their own chemical libraries and began to implement this technology as well through the 1990s when high throughput screening became popular.

As the technology became more widely available, its evolution and optimization progressed. Plates became miniaturized, while the number of wells increased, allowing for an even greater number of samples to be tested. Current screens can even be conducted with 1,536, 3456, well plates (Janzen, 2014 and Pareira; Williams, 2007; Wleklinski *et al.*, 2018). This was accomplished through the automation of the screen; by allowing robots to conduct and monitor the screens, the rate of the screen is no longer limited by human ability.

Process of High Throughput Screening

Target Identification and Validation

The first step in developing a drug through high throughput screening is identifying the target of the screen, meaning, the protein or cellular pathway whose behavior the researcher wishes to modify. This can often be based on decades worth of data into a particular protein's role in the cell, and the structure of the protein could have possibly been identified. However, as was pointed out by William P. Janzen, the former Director of Assay Development and Compound Profiling at University of North Carolina, "in reality, the more you know about a target, the less likely you are to undertake a screen, making it almost a tool of desperation in *de novo* discovery." (Janzen, 2014) Since high throughput screening is a method that does not require a complete understanding of the target, it is useful for proteins or targets that are not well understood. High throughput screening technology is more interested in finding a drug that modifies the target than understanding exactly how it accomplishes the change.

One useful tool for initially identifying a target for a potential drug is RNA interference. Double-stranded RNA that codes for the candidate protein target is introduced into cells and triggers cellular mechanisms that result in the degradation of cellular mRNA coding for the target. This prevents the mRNA from being translated, and the effect on the cell can be assessed. In this way, a particular gene's role in the cell can be elucidated. (Kroenke *et al.*, 2004)

Target identification can also be accomplished using clustered regularly interspaced short palindromic repeats (CRISPR) technology. When the CRISPR technology targets a gene, it cuts the DNA in a particular location, preventing proper expression of the target gene. This loss-of-function test can provide insight into how the cell reacts when the target is not present as usual. (Coussens *et al.*, 2017) While a target's role can be identified by observing what happens when it is not present, it can also be just as useful to see what occurs when it is overexpressed. Using CRISPR, cell lines have been created to overexpress certain genes, allowing researchers to identify the target's role in that way.

Assay Development

Once a target has been identified, the next step is to develop an assay that can be used to identify a hit molecule. A hit is a compound that produces the desired result in an initial screen. A hit can be turned into a lead compound if its effect is confirmed by additional assays, ones that show "biologically relevant activity that correlates to the target" (Fox *et al.*, 2006). The initial assay should be developed so that a simple readout identifies compounds that have the desired interaction with the target. Radioactivity, fluorescence, and absorbance have all been used in assays to provide data for target validation. For example, if the target is a protein that requires a ligand to bind for the cells to proliferate, the concentration of cells can be measured using spectrophotometry. Cells in which this crucial protein is functioning will survive and reproduce, thus having a higher concentration than samples in which the necessary protein did not function.

Targets can be exposed to small molecules in several different contexts. The first is a biochemical assay, which "utilizes a purified target protein of interest and measures the binding of ligands or the inhibition of enzymatic activity in vitro" (Blay *et al.*, 2020) In biochemical assays, the target is isolated from the cell and the effect of molecules is evaluated in that context. Typically, in a biochemical assay, the only effect measured is if the molecule bonds to the target at all. Then those hit molecules that have been identified as binding to the target can be evaluated further in additional assay. This technique is simple to conduct because it is performed without the confounding factors that can occur in a cellular context. However, many compounds identified in biochemical assays as hits fail to have similar effects once evaluated in the cell. Another issue that emerges when conducting screens independent of cells is permeability; even when a molecule can induce a particular effect on a target, it cannot reach it due to the cell membrane's selective permeability. Even more alarmingly, many of the hits can be found to be toxic to the cell. (An, 2010)

This failure of hits to develop into viable leads has caused researchers to pursue cell-based assays for high throughput screens. Cell-based assays place the target in the cellular context, to measure the ability of a molecule to induce a particular measurable result (Blay, 2020). By conducting the initial screen within the cellular setting, researchers can avoid many of the issues that cause the failure of hits from being developed into actual drug leads.

Recently, due to the high cost associated with performing high throughput screening, researchers have used computers to narrow down the chemical libraries used in the screens. A screen can be conducted *in silico*, virtually, using computer-aided drug discovery (CADD) tools.

A computer can screen the libraries for compounds that are likely to bond to the target based on either the structure of the target or the chemical's similarity to the target's ligand. Then only molecules that have been identified as potential hits will be evaluated in a physical assay. (Leelananda and Lindert, 2016) The benefit of using computer modeling is that it minimizes the financial cost of testing a full chemical library. However, conducting a screen *in silico* requires substantial knowledge of the target structure or its ligand's shape. Thus, it is not as useful in cases when the target is not as well-documented.

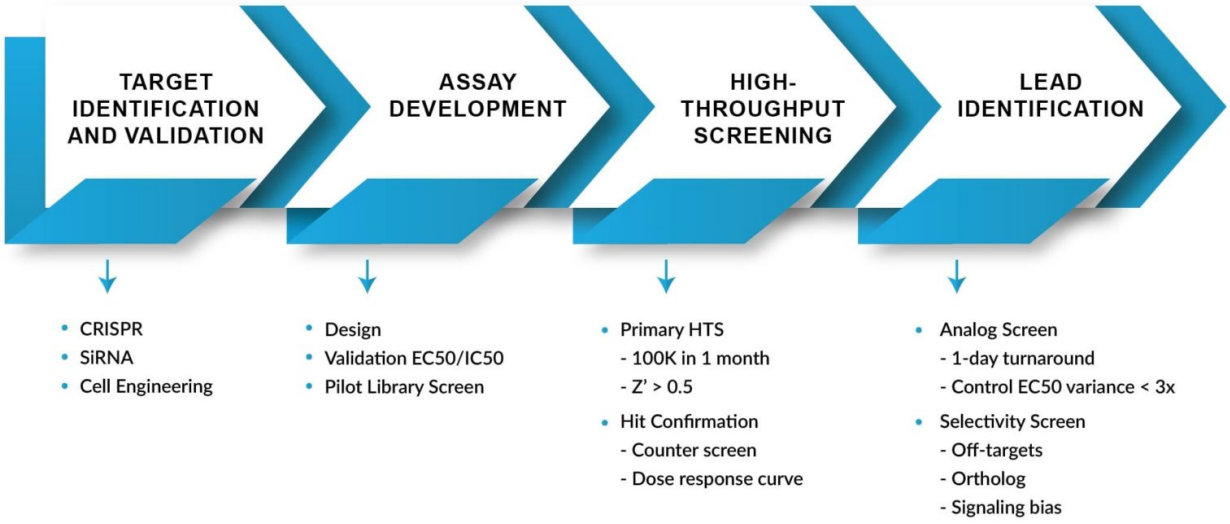

Figure 1: a representation of the multi-step process of high throughput screening (Multispan, n.d.)

SLC13A5 Deficiency and NaCT

Introduction

The sodium-citrate transport protein, NaCT, is encoded by the gene SLC13A5, found on chromosome 17. This transporter brings citrate into cells, where it can enter the Krebs Cycle, or be used for fatty acid synthesis. Citrate can also be produced intracellularly, but cells rely on the extracellular citrate for proper cellular functioning. SLC13A5 Deficiency is a rare autosomal recessive condition in which NaCT is misfolded, and citrate can't enter the cells. Subjects with this disorder experience early-onset epilepsy, developmental delays, and tooth dysplasia. Considering the severe nature of the disease, it is crucial to develop a drug to address the condition. (Klotz *et al.*, 2016; Hardies *et al.*, 2015; Thevenon *et al,* 2014) To date, there have only been about 50 recorded cases of SLC13A5 deficiency cases. (National Organization for Rare Disorders, 2017) SLC13A5 deficiency is a good target for a drug to be discovered by a high throughput screen since it is caused by a mutation in a single protein, so a molecule could be discovered that interacts with NaCT to recover protein function, and hopefully cure the disease.

Assay Development for High Throughput Screening for NaCT Activator

Ideally, a drug to treat SLC13A5 would be a small molecule that would interact with the mutated NaCT restoring proper function to the protein. This would allow citrate to enter cells in appropriate concentrations, and mitigate the symptoms of the disease. This setup is a perfect candidate for high throughput screening, in which the target protein would be exposed to thousands of small molecules, and the one that causes NaCT to function would be identified as a hit and developed further into a drug lead.

In the summer of 2019, the author conducted research in the lab of Da-Neng Wang at the Skirball Institute of Biomolecular Medicine at NYU Langone Hospital to develop an assay to be used in a screen to recover NaCT function. The assay method chosen for this project was a cell-based assay, in which the protein of interest was evaluated in a cellular context. SLC13A5 was transformed into *E. coli* and overexpressed. *E. coli* was chosen as the host because the bacteria do not express a native citrate transporter under aerobic conditions. Once it was confirmed that NaCT was successfully transformed in the cells by DNA sequencing, *E. coli* were grown in minimal media, in which citrate was the singular carbon source. The only way for the cells to reproduce was if they were able to uptake citrate from the growth media, which requires that NaCT folded and functioned correctly. (Figure 2)

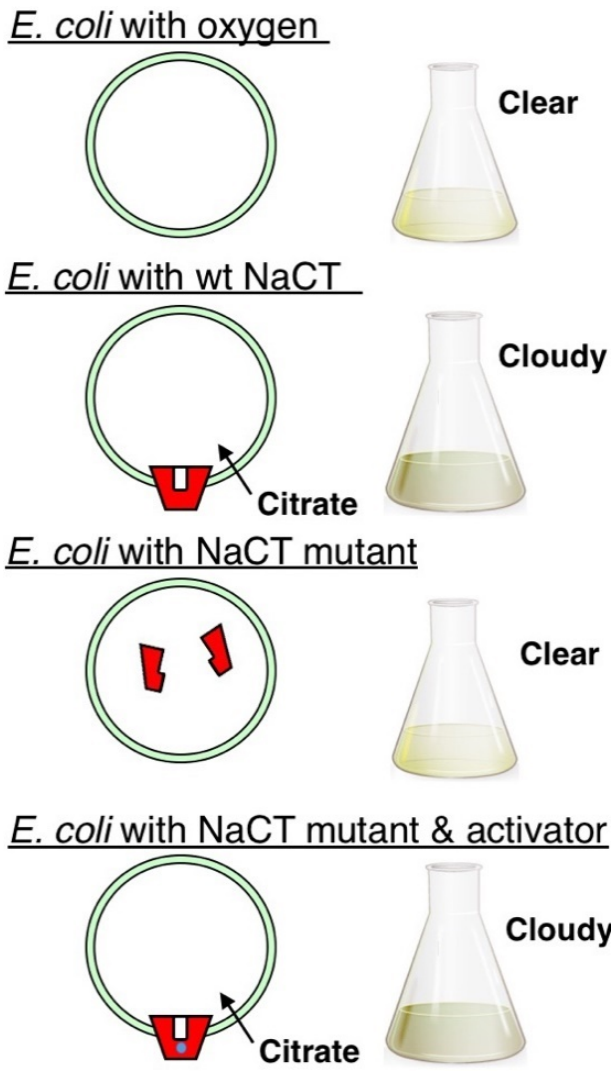

Figure 2: Steps of screening for activators of NaCT mutants in *E. coli* cells in minimal media with citrate as the sole carbon source under aerobic conditions.

Creating a Construct for Expressing NaCT in *E. coli*

Expressing a human protein in bacterial cells has many challenges, and membrane proteins are particularly difficult to work with because of the large number of hydrophobic regions. Due to the transcription, translation, and packaging mechanisms being so different in

bacterial cells, the gene cannot simply be inserted into the bacterial cell's genome. Instead, to aid with proper folding and localization of the NaCT, fusion proteins were added to the protein of interest. Fusion proteins are small, globular, hydrophilic proteins attached to proteins to aid in their folding and localization in cells. Several different proteins have been proven to be successful in the overexpression of membrane proteins in *E. coli* cells. The ones chosen for this experiment were YaiN, and YbeL, identified for simplicity as α and β respectively, and Green Fluorescent Protein (GFP). YaiN is part of an operon *in E. coli* likely involved in the breakdown of formaldehyde, and the function of YbeL is unknown. These proteins have all been used to overexpress membrane proteins in *E. coli* in different combinations. (Leviatan *et al.,* 2010; Drew *et al.*, 2006). GFP has the added benefit of providing a way to visualize the protein in the cell and measure the amount by fluorescence. (Kain, 1999)

Unfortunately, it is impossible to predict which fusion protein combination will be effective to overexpress a particular protein of interest. Thus, the early stage of this assay development depended on guessing the right construct. The fusion proteins can be added on the protein of interest to its N terminus, C terminus, or both. Before attempting to express the mutated NaCT, first, it was necessary to express the wild-type NaCT to identify the correct construct of fusion proteins. Once a combination was identified that allows NaCT to be expressed properly in *E. coli*, the same construct could be created for a mutated NaCT. (Figure 3)

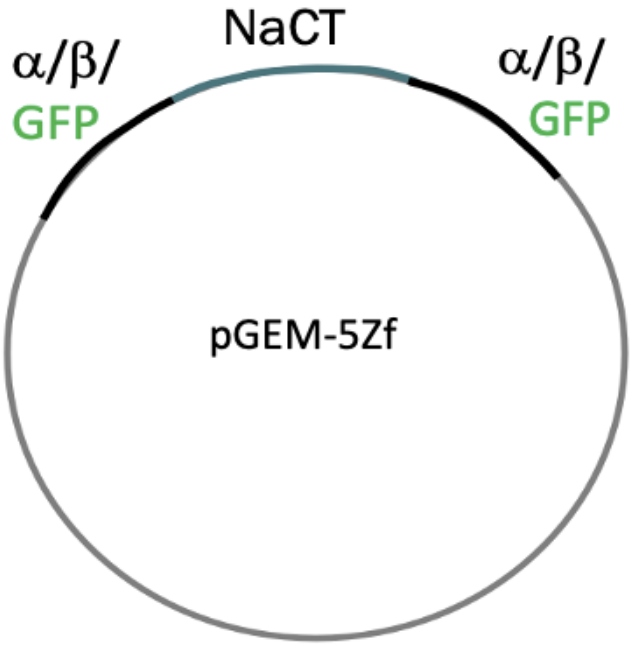

Figure 3: A representation of NaCT cloned with fusion proteins α, β, and GFP into the complementation vector, pGEM-5Zf

The sequence of wild-type NaCT was amplified by PCR and then restriction cloned into pET vectors on which the fusion proteins were located. Then the entire fusion protein-NaCT constructs were excised from the pET vectors and cloned into a complementation vector, pGEM-5Zf, which allows for the unregulated expression of NaCT. These vectors were transformed into *E. coli* grown in minimal media with citrate as the sole carbon source. Cells that grew successfully contained NaCT that was being expressed sufficiently and functioning properly, with the correct fusion protein construct.

Creating a Positive Control Using CitS

To evaluate whether cells grew in the proper levels in the carbon-poor environment, growth had to be compared to a positive control. CitS, a citrate transporter from *K. pneumoniae* had previously been successfully expressed in *E. coli* (van der Rest, 1992). When CitS is expressed in *E. coli* and grown in a carbon-poor environment, *E. coli* grows in a normal fashion when compared to the negative control, cells without the added citrate transporter. (Figure 4) This validates the hypothesis that *E. coli* is only able to grow in aerobic environments with citrate as the only carbon source if a foreign citrate transporter is added. Cells that express CitS act as excellent positive controls for ones in which NaCT is being cloned since if NaCT is being expressed properly, the cells would grow at the same levels.

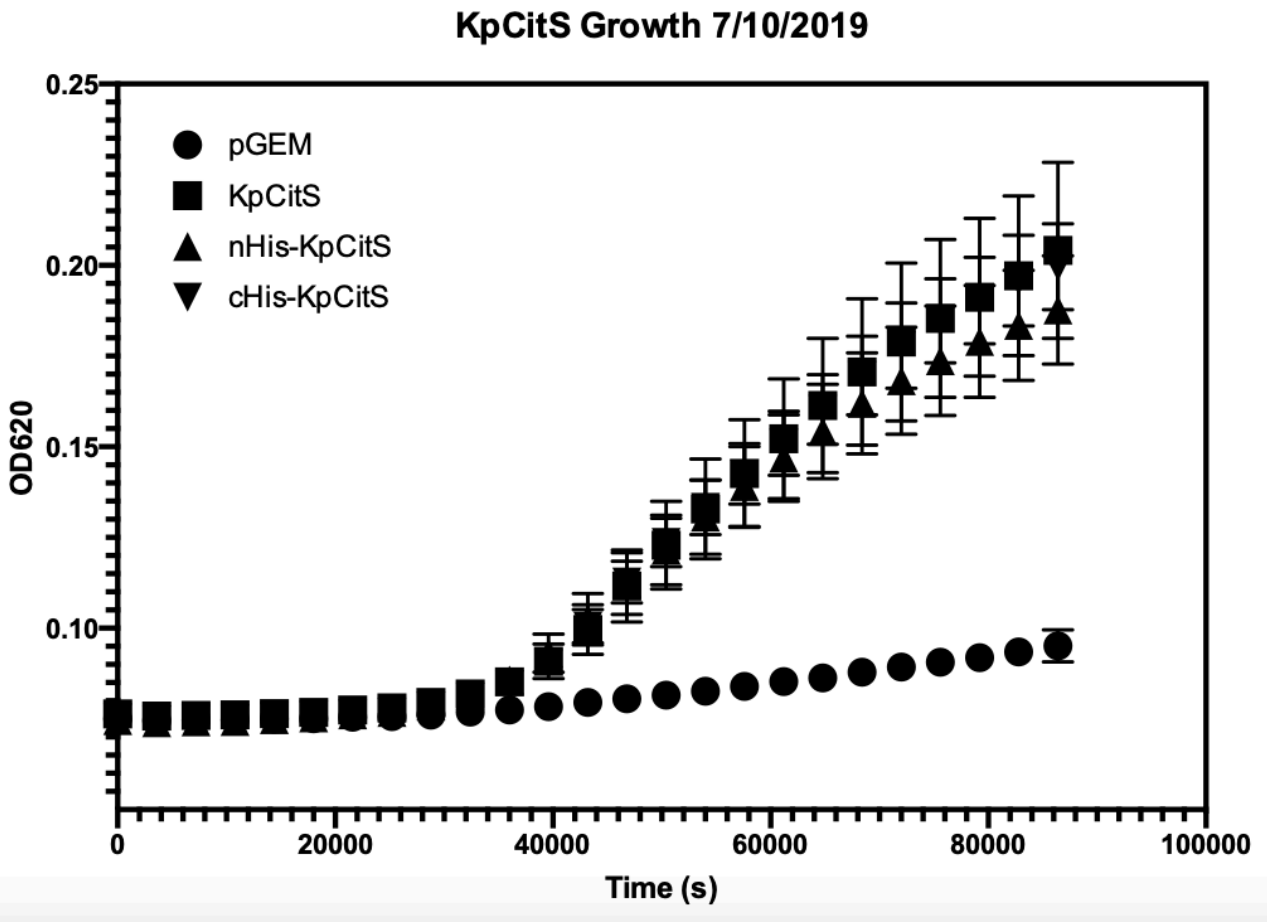

Figure 4: Growth of *E. coli* with CitS, with the addition of a poly-histidine tag to the N or C terminus of CitS compared to cells containing an empty vector (pGEM) without CitS. The poly-histidine tag was intended to allow for protein visualization on a western blot.

Many different combinations of fusions proteins and NaCT were attempted, but to date, none have been successful in NaCT in *E. coli*. The combinations tried so far have been α- NaCT, β-NaCT, α-NaCT-β, and β-NaCT-β. Bacterial proliferation was measured by absorbance at 620 nm. When compared to the growth expressed by the cells with the CitS protein, the cells with α-NaCT and β-NaCT did not exhibit growth that would indicate that the NaCT was functioning

properly. The cells displayed absorbance that was more similar to, if not worse than, *E. coli* without any citrate transporter added, ones with an added empty vector. (Figure 5)

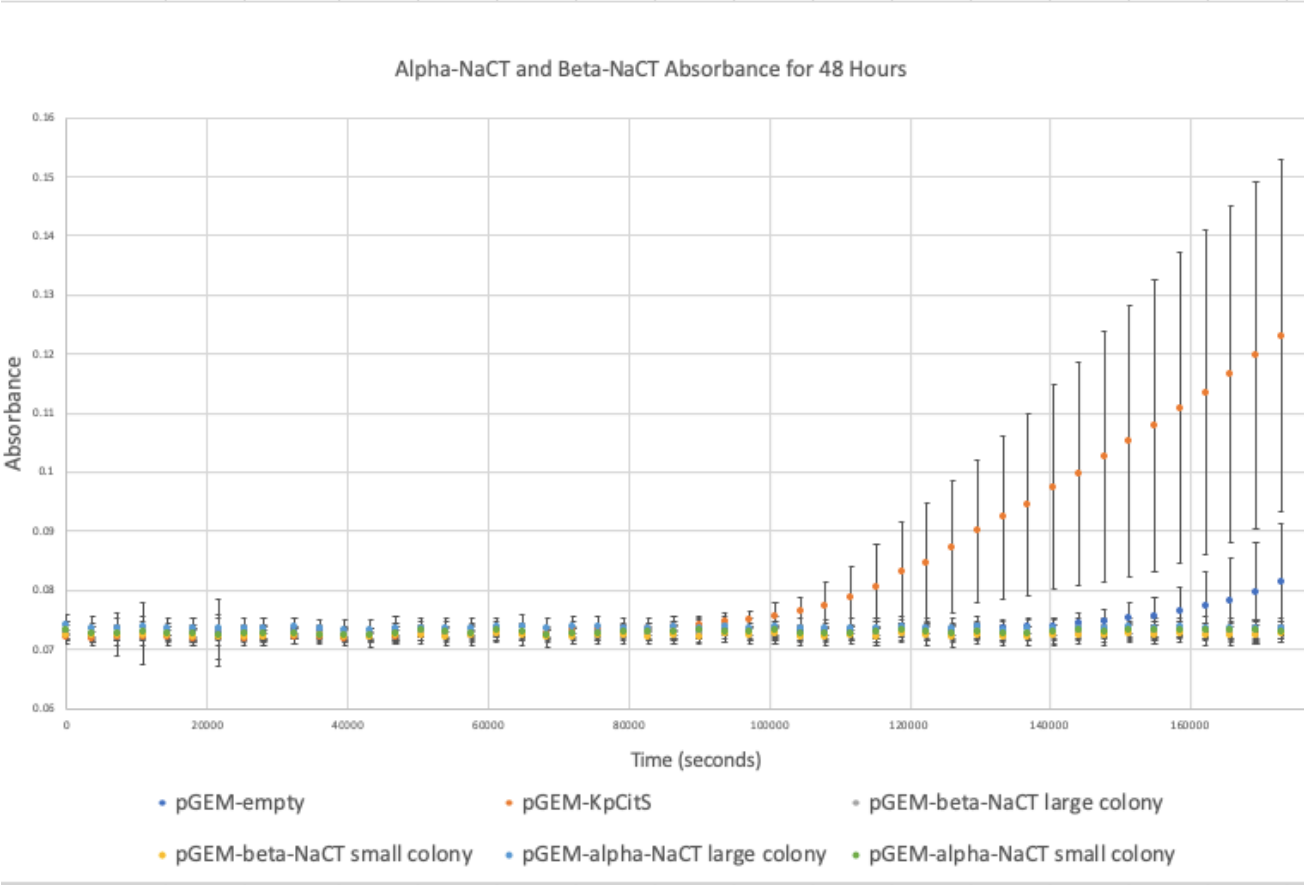

Figure 5: Growth of *E. coli* with the addition of the α-NaCT, or β-NaCT constructs in the complementation vector. Two different mutant colonies were selected from each construct and grown in minimal media.

Expressing NaCT on pET Vectors with the Addition of IPTG

The experiment protocol was adjusted in response to difficulties in ligating the fusion protein-NaCT constructs with the complementation vector. When the complementation vector was transformed into *E. coli*, the cultures displayed significantly stunted growth even in a

carbon-rich environment when compared to cells without the vectors. While this may be due to human error, it was also hypothesized that the NaCT was being overexpressed and becoming toxic to the cells. This could be due to an abundance of misfolded protein aggregating in cells resulting in death, a phenomenon that is a feature of many neurological disorders. (Chen, 2011) Alternatively, NaCT could be overexpressed in the cells to an extent that it was introducing an excess of citrate to a toxic level.

Whatever the reason for the failure in expressing NaCT using the complementation vector, an alternative solution was attempted that excluded that step in the protocol. Instead, *E. coli* was grown with the addition of the NaCT in the pET vector on which the fusion proteins were found. The solution had initially been rejected for several reasons. The first is that the goal was to express NaCT in as high of a concentration as possible, to maximize success using the complementation vector. However, the problem may have been due to overexpression, and transcription from the pET vector would lead to a more modest amount. Additionally, the location on the pET vector upon which the fusion proteins were located provided a unique challenge as it was the lac operon. To induce the expression of the proteins added to the vector, lactose would have to be added. However, introducing lactose into the carbon-poor media would introduce a new carbon source, thus making it impossible to determine if the cells were actually utilizing citrate or not. The solution to this issue relied on the use of isopropyl-β-D-thiogalactoside (IPTG), a lac operon inducer that cannot be modified by the cell to be used as a carbon source. (Donovan *et al.,* 1996) With the addition of IPTG, NaCT can be expressed on the pET vector without negating the purpose of the assay.

Using this new experimental setup, the fusion protein-NaCT constructs were transformed into *E. coli* and grown in carbon-poor media with the addition of IPTG. The constructs tested were α- NaCT, β-NaCT, α-NaCT-β, and β-NaCT-β. The results of this experiment were inconclusive since all of the cultures exhibited similar growth, even those without an added citrate transporter. (figures 5A and 5B) A western blot was performed to detect NaCT production in the samples, using the horseradish peroxidase (HRP) antibody, which binds to the fusion proteins. A faint band was found in all of the samples at around 17 kDa, but no NaCT was detected at 65 kDa. (Figure 6) It is still unclear why those cells all grew despite the lack of carbon in the media. It was likely due to human error or a mistake in the experimental setup.

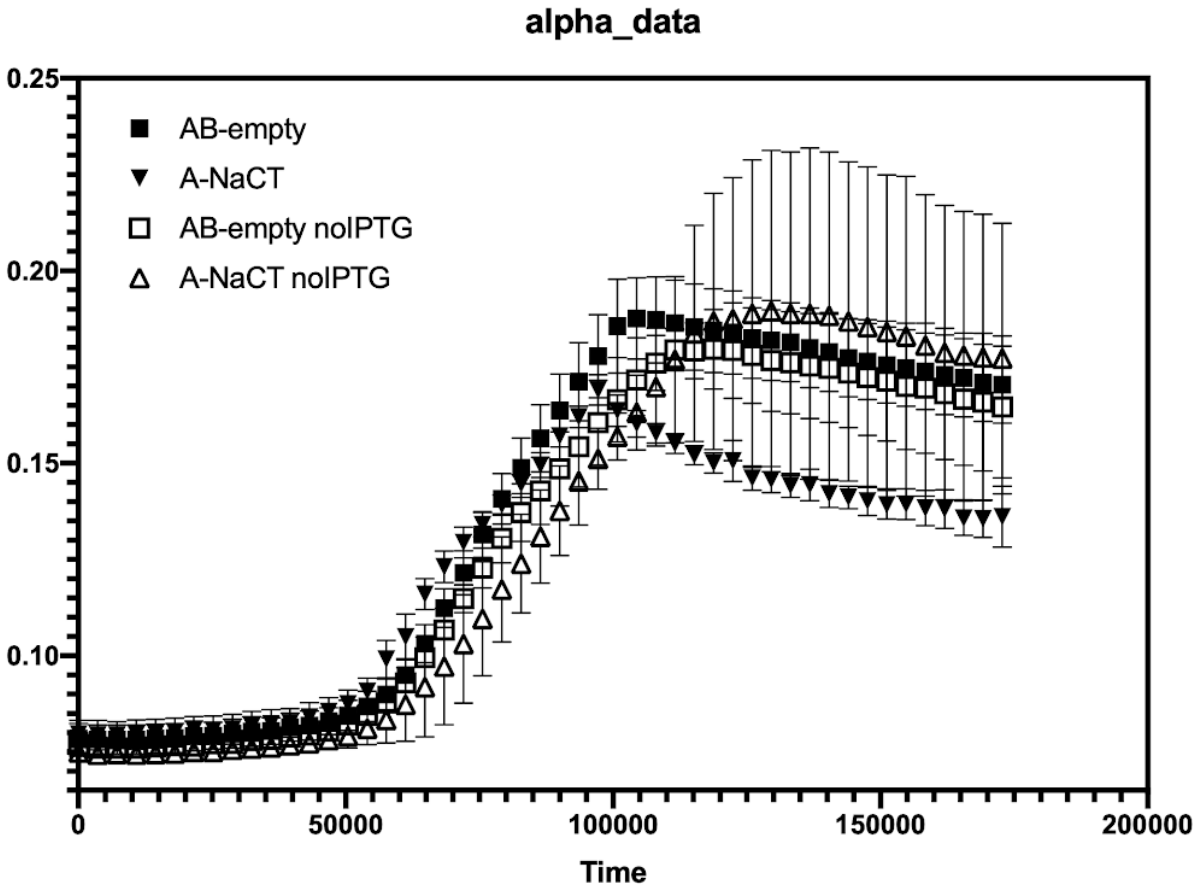

Figure 5A: Growth of *E. coli* with NaCT, using the αNaCT combination on the pET vector, with the addition of IPTG. Growth was compared to cells containing empty vectors, and cells without the addition of IPTG.

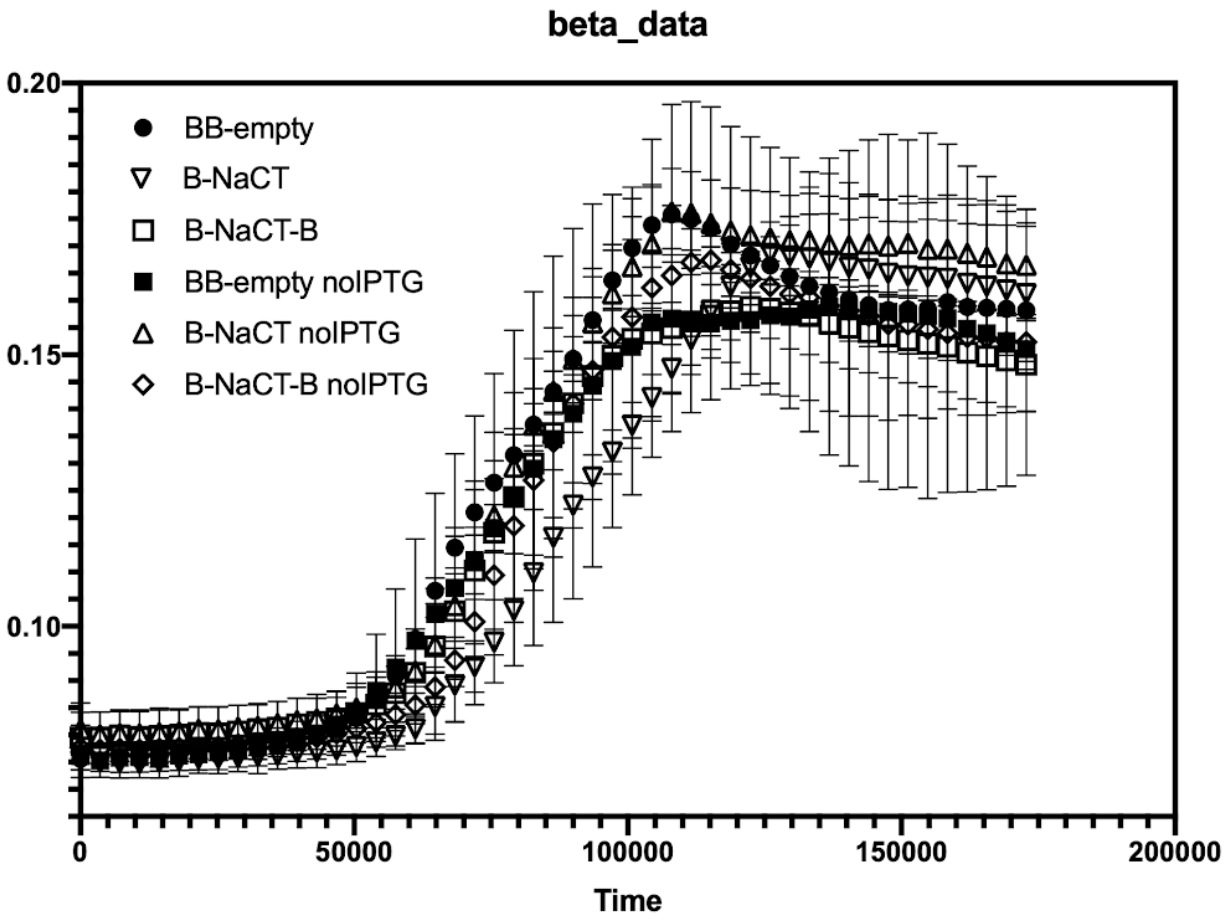

Figure 5B: Growth of *E. coli* with NaCT, using the βNaCTβ, and βNaCT combinations on the pET vector, with the addition of IPTG. Growth was compared to cells containing empty vectors, and cells without the addition of IPTG.

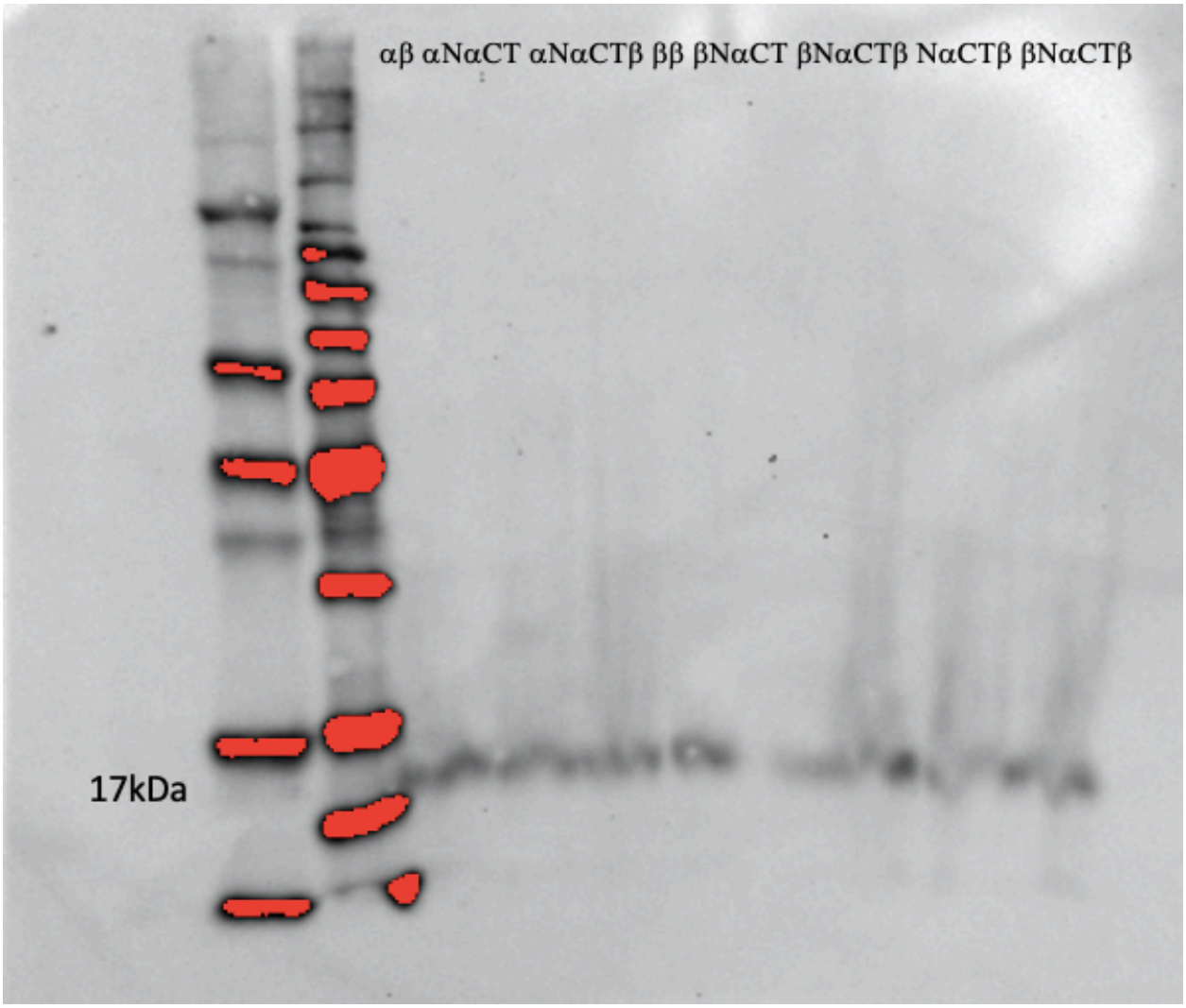

Figure 6: A Western Blot of proteins isolated from *E. coli* transformed with different fusion protein-NaCT complexes on the pET vector.

Conducting a Screen

Although a functioning assay has not yet been created, hopefully with continued optimization, one will be successful. The next stage of the assay optimization would be to express the mutated form of the NaCT with the same combination of fusion proteins. This would allow the protein to fold in the same way that it does in human cells. The cells would again be

grown in carbon-poor media, and would not be viable, due to their mutated transporter. The screen would then test the large chemical libraries to evaluate whether any of the small molecules induce proper NaCT function.

Recent Developments

Recently, the Wang lab at NYU Skirball was successful at solving the structure of NaCT. This was accomplished through cryo-electron microscopy. With a greater understanding of the structure of the protein, more insight was gained into the different classes of NaCT mutations, information that can be valuable in identifying an activator molecule. (Sauer *et al.*, 2021) Through solving the structure of NaCT, the authors were able to divide protein mutations into two categories. Type I mutations affect the localization of the protein and make it susceptible to proteolytic degradation. On the other hand, while type II mutated proteins can reach the cell membrane, the mutations prevent proper binding and transport of the substrates. These types can be further subdivided into groups based on the exact class and location of the mutation. This greater understanding of the causes of the mutations makes it clear that each type would require unique treatment since the mutations exist in such a range of locations in the protein. Now that the structure of NaCT and its mutants is known, perhaps an *in silico* assay would be an effective next step in identifying a small molecule that could bind to the transporter and induce proper functioning.

Arguments Against High Throughput Screening

As evident from the example of NaCT, developing a workable assay for high throughput screening is challenging and time-consuming. Often, a potential screen will fail in the development stages, even before testing the chemical libraries. While the actual high throughput screen can be conducted in as short as a day, setting it up can take years. Investing years into developing an essay only for it to fail to yield viable drug leads can be frustrating and wasteful.

Other objections have been raised again the practice of high throughput screening. The first is the high cost associated with the process, in particular the initial investment in automated equipment. But even after the original investment, the cost of running a screen is expensive. At the University of Wisconsin Carbone Cancer Center, which runs its own high throughput laboratory and provides services for those outside the university, a screen can cost anywhere from ten cents to one dollar per well. (University of Wisconsin Carbone Cancer Center, n.d.) In a modest screen that tests 10,000 compounds, this cost may be manageable, but most screens exceed that number significantly, perhaps testing half a million compounds. This cost may be insignificant in the pharmaceutical industry, whose top 10 companies spent a collective $83 billion on research and development in 2019. (Congressional Budget Office, 2021) But a university lab, whose funding relies on grant money, will likely find that cost too high to be a worthwhile endeavor.

Many object to high throughput screening on an intellectual level, arguing that it stifles creativity and intelligence in the pharmaceutical industry. Roger Lahana, the Vice President for the French biopharmaceutical company Synt:em, wrote that high throughput screening is

"irrational" and argued that "screening the Sahara desert to get one given atom of one given grain of sand would be a small task in comparison to screening all the potential drug-like (not to mention lead-like) compounds to identify every possible drug" (Lahana, 2003) In the Harvard Business Review, Jean-Pierre Garnier, the CEO of GlaxoSmithKline from 2000-2008, argued that the way to improve productivity in the pharmaceutical industry is to invest in intelligent drug design and create smaller groups of researchers. He explained that while sequencing the genome and high throughput screening could reflect great strides in research development, "tools themselves are no substitute for first-rate scientific minds; a fool with a tool is still a fool." (Garnier, 2008) Many argue that by shifting much of the initial drug discovery to high throughput screening, pharmaceutical companies are becoming less creative, and attempting to develop drugs through brute force, instead of intelligent design.

Arguments In Favor of High Throughput Screening

While there are many criticisms of high throughput screening, there are many benefits as well. One of the greatest is that it provides a way to develop drugs for lesser-known targets. It is not necessary to have solved a protein's structure or understand an entire biological pathway to perform a screen to modulate the target. In that way, high throughput screening is a valuable tool for more elusive and not well-understood targets. Creating a truly novel drug is rare and relies on having a complete understanding of the target. Instead, utilizing high throughput screening can allow for the creation of drugs for targets whose structures have not been solved. This is particularly helpful for targets that are proving difficult to study, or to which sufficient time and resources have not been dedicated. These less popular targets are excellent subjects for high throughput screens. Another

benefit of high thoughput screening is that it provides a standard route of drug development. Instead of stumbling around waiting to get lucky, researchers can follow an established pathway to reach success. (Figure 1)

While the exact success rate of high throughput screening has been a matter of debate, it certainly has yielded positive results. In a 2001 study that surveyed directors of high throughput directors and suppliers, it was reported that 43% of targets generated leads from screens. (Fox S. *et al.,* 2006) More recent data from 2011 put the success rate at over 60%. In an analysis of 60 drugs approved from 1978-2008, nearly a quarter could be attributed to high throughput screening. (Perola, 2010) High throughput screening is being utilized and having success in developing drugs that are being brought to the market.

High Throughput Screening in Academia

Initially, it was only pharmaceutical companies engaging in high throughput screening, but lately, there has been an increased interest in the technology from academics. In the past three decades, the utilization of this technology has expanded outside of the pharmaceutical industry and is now performed in a select number of universities. Universities were drawn to high throughput technology for the same reason as companies: the hopes that it would hasten the research process and provide an easier way to identify drugs. However, in academia, there are still “unrealistic expectations on the part of the principal investigator sponsoring the target and access to high-quality libraries.” (Janzen, 2014) When academics started to participate in high throughput screening, there was a perception that it would provide an easy route to publishable

results, but the harsh realities of drug development still exist within academia. 20% of screens do not result in any potential hits for drug development, and another 25% of those hits fail to develop into drug leads. (Paul *et al.*, 2010) In the case of academia, money is usually fought for in competitive grant applications, and to have no data to show for three years of effort is incredibly disappointing. For a single research team to put all their time and hard-earned money chasing potentially nothing may not be useful.

The pharmaceutical industry seems to be better suited to utilizing high throughput screening technology from both a cost and return on investment standpoint. In a study evaluating the impact of high throughput screening on the oncology drug field, the authors compiled a list of fifteen cancer drugs approved from 2001-2015 that were identified from a small molecule screen. (Coussens, *et al,* 2017). In every case, one of the 35 largest biomedical companies in the world by revenue was involved in the development of the drug. The initial investment in the automated equipment necessary to perform high throughput screens is incredibly expensive and may be beyond the means of academia. The price of equipment varies greatly depending on the specifics of the machines. In a 2007 study that interviewed high throughput screening laboratories, one director reported that the price of some instruments can be a quarter to half a million dollars. (Downey *et al.,* 2007) This price has likely increased significantly in the interim and does not include the price of curating and storing a robust chemical library, which can reach 500,000 compounds (Macarron *et al,* 2011) Academia's creativity and budgets would be put to better use in more creative and rational drug design, that utilizes the intellectual capability of its researchers.

Conclusions

High throughput screening has been adopted by pharmaceutical companies and academics alike as a tool for drug development. It has been successful in developing novel drugs since its rise in popularity in the 1990s. The process of target identification and assay development can be lengthy and disappointing, as demonstrated through the NaCT project. However, it does often yields positive results, and is particularly valuable for targets that are not well-known. This standardized method of drug development has been largely accepted, despite the high costs associated with utilizing it. It has been largely profitable in the pharmaceutical industry and several drugs are currently in use that were discovered using high throughput screening. Many institutions put money into high throughput screening technology in the hopes that they would achieve similar success. However, due to the high initial investment and the continuous cost of maintenance, it is not a justifiable endeavor for academia.

Acknowledgments

I want to thank the S. Daniel Abraham Honors Program and Dr. Wachtell for my amazing experience at Stern College for Women, and for the incredible learning opportunities made available to me. I would also like to thank my mentor, Dr. Schuck, for her continued guidance and feedback. I want to express my overwhelming appreciation to Dr. Da-Neng Wang at the Skirball Institute at NYU for welcoming me into his lab, and Dr. David Sauer for guiding and encouraging me through the project. Finally, I would like to thank my parents for their continued love and support.

Works Cited

American Chemical Society International Historic Chemical Landmarks, n.d. Discovery and Development of
Penicillin. http://www.acs.org/content/acs/en/education/whatischemistry/landmarks/flemingpenicillin.html (retrieved April 16, 2021).

MedlinePlus, n.d. Penicillin V Potassium: MedlinePlus Drug Information.
https://medlineplus.gov/druginfo/meds/a685015.html (retrieved April 16, 2021)

Foundation for a Drug Free World, n.d. Painkillers: A Short History.
https://www.drugfreeworld.org/drugfacts/painkillers/a-short-history.html. (retrieved April 16, 2021)

Mavromoustakos T. *et al.,* Strategies in the rational drug design. Curr Med Chem. 2011;18(17):2517-30.

Arulsamy A, Kumari Y, Shaikh M.F., Reverse pharmacology: fast track path of drug discovery. Pharm Pharmacol Int *J*. 2016;4(3):358-359.
Macarron R. *et al.,* Impact of high-throughput screening in biomedical research. Nat Rev Drug Discov. 2011 Mar;10(3):188-95.

Attene-Ramos, M.S., Austin, C.P., Xia, M., High Throughput Screening; Elsevier: Amsterdam, The Netherlands, 2014; pp. 916–917.

Janzen W.P., Screening technologies for small molecule discovery: the state of the art. Chem Biol. 2014 Sep 18;21(9):1162-70.

Pereira D.A., Williams J.A., Origin and evolution of high throughput screening. Br J Pharmacol. 2007 Sep;152(1):53-61.

Wleklinski M. *et al.*, High throughput reaction screening using desorption electrospray ionization mass spectrometry. Chem. Sci*.,* 2018, 9**,** 1647-1653

Kroenke, A. *et al.*, Drug Discovery World (DDW), April 3, 2004, Using RNAi to Identify and Validate Novel Drug Targets: Targeting Human Kinases with an siRNA Library.
https://www.ddw-online.com/using-rnai-to-identify-and-validate-novel-drug-targets-targeting-human-kinases-with-an-sirna-library-1743-200404/ (retrieved April 17, 2021)

Coussens NP *et al.,* Small-Molecule Screens: A Gateway to Cancer Therapeutic Agents with Case Studies of Food and Drug Administration-Approved Drugs. Pharmacol Rev. 2017 Oct;69(4):479-496.

Fox S. *et al.,* High-throughput screening: update on practices and success. J Biomol Screen. 2006 Oct;11(7):864-9.

Blay V. *et al.*, High-Throughput Screening: today's biochemical and cell-based approaches, Drug Discovery Today, Volume 25, Issue 10, 2020,Pages 1807-1821

An WF, Tolliday N. Cell-based assays for high-throughput screening. Mol Biotechnol. 2010 Jun;45(2):180-6.

Leelananda SP and Lindert S., Computational methods in drug discovery. Beilstein J Org Chem. 2016 Dec 12;12:2694-2718.

Multispan, n.d. The Power of Integrating Screening and Assay Development. https://multispaninc.com/articles/high-throughput-screening-done-right/ (retrieved April 19, 2021)

Klotz J *et al*., Mutations in the Na(+)/citrate cotransporter NaCT (SLC13A5) in pediatric patients with epilepsy and developmental delay. Mol Med. 2016 May 26;22:310–21.

Hardies K *et al.*, Recessive mutations in SLC13A5 result in a loss of citrate transport and cause neonatal epilepsy, developmental delay and teeth hypoplasia. Brain. 2015 Nov;138(Pt 11):3238-50.

Thevenon J *et al.,* Mutations in SLC13A5 cause autosomal-recessive epileptic encephalopathy with seizure onset in the first days of life. Am J Hum Genet. 2014 Jul 3;95(1):113-20.

National Organization for Rare Disorders, 13 December 2017, SLC13A5 Epileptic Encephalopathy. https://rarediseases.org/rare-diseases/slc13a5-epileptic-encephalopathy/

Leviatan S. *et al.,* Combinatorial method for overexpression of membrane proteins in Escherichia coli. J Biol Chem. 2010 Jul 30;285(31):23548-56. Epub 2010 Jun 4.

Drew, D., *et al.* Optimization of membrane protein overexpression and purification using GFP fusions. *Nat Methods* 3, 303–313 (2006).

Kain SR. Green fluorescent protein (GFP): applications in cell-based assays for drug discovery. Drug Discov Today. 1999 Jul;4(7):304-312.

van der Rest ME, Molenaar D, Konings WN. Mechanism of Na(+)-dependent citrate transport in Klebsiella pneumoniae. J Bacteriol. 1992 Aug;174(15):4893-8.

Chen B *et al.*, Cellular strategies of protein quality control. Cold Spring Harb Perspect Biol. 2011 Aug 1;3(8):a004374.

Donovan, R.S., Robinson, C.W. & Glick, B.R. Review: Optimizing inducer and culture conditions for expression of foreign proteins under the control of the *lac* promoter. *Journal of Industrial Microbiology* 16, 145–154 (1996).

Sauer, D.B. *et al.* Structure and inhibition mechanism of the human citrate transporter NaCT. *Nature* 591, 157–161 (2021). https://doi.org/10.1038/s41586-021-03230-x

University of Wisconsin Carbone Cancer Center, n.d., Services, Equipment & Pricing, https://cancer.wisc.edu/research/resources/ddc/smsf/equipment-services/#:~:text=Generally%20HTS%20assays%20cost%20from,prior%20to%20beginning%20a%20project (retrieved April 18, 2021)

Congressional Budget Office, Research and Development in the Pharmaceutical Industry, April 2021

Lahana R., Who wants to be irrational? Drug Discov Today. 2003 Aug 1;8(15):655-6.

Garnier JP, Rebuilding the R&D Engine in Big Pharma, Harvard Business Review. 2008 May, https://hbr.org/2008/05/rebuilding-the-rd-engine-in-big-pharma

Paul, S. *et al.* How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* 9, 203–214 (2010)

Perola E, An Analysis of the Binding Efficiencies of Drugs and Their Leads in Successful Drug Discovery Programs. Journal of Medicinal Chemistry 2010 *53* (7), 2986-2997

Downey W, Nicely H, Sopchak L. Drug Discovery World (DDW), December 27, 2007, Drug Discovery World, Needs and Challenges in High Throughput Screening, https://www.ddw-online.com/needs-and-challenges-in-high-throughput-screening-554-200712/. (retrieved April 21, 2020)