

Commentary

Problem choice and decision trees in science and engineering

Michael A. Fischbach^{1,2,3,*}¹ChEM-H Institute, Stanford University, Stanford, CA 94305, USA²Department of Bioengineering, Stanford University, Stanford, CA 94305, USA³Chan Zuckerberg Biohub, San Francisco, CA 94158, USA*Correspondence: fischbach@fischbachgroup.org<https://doi.org/10.1016/j.cell.2024.03.012>

Scientists and engineers often spend days choosing a problem and years solving it. This imbalance limits impact. Here, we offer a framework for problem choice: prompts for ideation, guidelines for evaluating impact and likelihood of success, the importance of fixing one parameter at a time, and opportunities afforded by failure.

There are a finite number of weeks in one's career. While scientists and engineers generally acknowledge that time is precious and we should strive to use it for greatest impact, agreeing with this and putting it into action are two different things. This is particularly challenging when choosing which problem to work on—a decision that can impact how you spend several years of your life.

Having observed that graduate students are taught almost everything about the theory and practice of science except how to pick a problem, and inspired by a previous piece on the topic,¹ we started a new class on problem choice in 2019 that is intended for first- and second-year graduate students. It has stimulated a great deal of discussion and spawned a similar course for new principal investigators at Stanford. What follows is a description of the content and potential impact of the course.

Spend more time on problem choice

A typical project for an incoming graduate student might involve 1–2 weeks of planning and 2–5 years of execution (Figure 1A). Once you choose a project, you are confined to a relatively narrow band of impact (Figure 1B); barring an unexpected surprise, the solution to a mediocre problem will have incremental impact, whereas solving an important problem will have greater impact. Even if you execute well, it is hard to make the solution to a middling problem interesting. In a very real sense, the problem you choose will influence the impact of

your work just as much as the quality of your execution.

In endeavors like writing software, the cycle time is short. You can just try it and see if it works; if not, you have only lost a couple weeks. This doesn't work in biology, chemistry, physics, and (non-software) engineering, where a typical project takes months to reach a go/no-go threshold and years to complete. A poor project chosen in haste can be hard to shed; inertia takes over and the sunk-cost fallacy is difficult to avoid.

It can help to reverse the polarity of our relationship with new ideas. We often treat them with reverence; before long, confirmation bias sets in and we begin looking for reasons why they will succeed and ignoring evidence suggesting they might fail. A common failure mode is to jump on the first idea and get started—this is probably the worst thing you can do. It is better to treat ideas like leeches that are trying to make a meal of your time. Consider their high points but treat them with skepticism, look for their warts, and take enough time to develop and evaluate several of them in parallel—comparison shopping leads to better decision making.

Once you choose a project, the landscape will change: you'll learn more, others will too, and technology develops. Critical thinking, evaluation, and course correction will continue throughout the project (see the “altitude dance” below), but it is better to take enough time to pick wisely in the first place.

Exercise “intuition pumps” and avoid common traps

There is no single way to generate new ideas, but certain prompts can help jumpstart the ideation process, and there are common traps to avoid (Table 1). A useful oversimplification is that most projects in the biological sciences involve perturbing a system, measuring it, and then analyzing the data (Figure 1C). New ideas generally fall under one of these headings; in each case, they involve an advance in technology or logic. Developing a new base editor or a method for constructing whole-genome CRISPR libraries is perturbation technology, whereas the use of an existing technology (e.g., a base editor) to perform deep mutational scanning is perturbation logic. A new tissue-clearing technique is measurement technology, while the use of tissue clearing to study liver fibrosis is measurement logic. Computation falls into multiple categories: analysis of cryo-electron microscopy or single-cell transcriptomic data could be considered measurement, whereas protein structure prediction and evolutionary theory warrant their own category—in which efforts can still be organized into technology (building a new algorithm or model) and logic (using it to make a discovery). As with any taxonomy, this one is imperfect but useful—it can help you learn, e.g., that your niche is perturbation logic or measurement technology.

What follows are a list of intuition pumps² to jumpstart the ideation process. In addition to considering your level of

interest, you might think about whether you would have a competitive advantage in carrying out the project. Perhaps you are an expert in a unique pair of skills (e.g., machine learning and cardiovascular development), or you have access to an unpublished data set that you can use to generate hypotheses (e.g., for genes linked to a disease phenotype).

Don't avoid risk; befriend it

A useful starting point for evaluating a new idea is to place it on a graph with two axes (Figure 1D): how likely it is to work vs. how much impact it will have. We will consider these one at a time.

To score a project's likelihood of success, we suggest a concrete exercise for trainees: an "assumption analysis" (Box 1). The idea is straightforward: for the project in question, list every assumption you are making from its inception through its conclusion. Assign two scores to each one: likelihood of success and duration of effort. Then look critically at the list. A project with a high-risk assumption that will not read out for >2 years is problematic, and one that will require multiple miracles to succeed should be avoided or refined.

The example we use in class (Box 1) outlines an effort to use single-cell RNA sequencing to identify new enteroendocrine cell types in the intestinal epithelium, an idea that is outdated but still illustrative. It becomes apparent through the analysis that there are two high-risk assumptions that do not read out until two years have passed: the existence of 2–3 new cell types can be validated experimentally, and selectively deleting one of these cell types in a mouse will yield a biologically relevant phenotype. This kind of risk profile is logistically

challenging and stressful. There are a few ways to fix it. For example, instead of identifying new enteroendocrine cells, the project could be reframed as an effort to learn more about enteroendocrine cell types that were previously known but incompletely characterized, like enterochromaffin cells. Or the analysis could focus instead on the liver, where easier options exist for genetic manipulation to validate new cell types. Or it could be

coupled to spatial transcriptomics to validate the existence of new cell types and help place them in a biogeographic context.

Importantly, the idea here is not to eliminate risk—risk-free projects tend to be incremental. Instead, the goal is to name, quantify, and work steadily to chip away at risk. As a corollary, when presenting an idea for a new project or startup company, be candid about risk—it has the paradoxical effect of making your case more convincing.

In some cases, it is possible to design a project that can succeed no matter how the data turn out. This type of project has the benefit that you won't feel inclined to root for one outcome over another. One common way of doing this is to characterize multiple candidates rather than a single one. Don't perform a genetic screen with one kinase or phosphatase, test a panel of them in parallel. Don't build one engineered bacterium, or adeno-associated virus, or lentivirus: make a library and test a pool of them. The likelihood of a successful outcome is greater, and the array of data showing how genetic or biochemical structure map to function will yield a richer model.

However, this is not always possible. If so, an important rule of thumb is to perform the go/no-go experiment at the earliest feasible moment.

This is true even if it requires some compromise; build a clunky prototype and see if it works, even a little.

Pick the right optimization function

The y axis of our idea evaluation graph—*How much impact will it have if it succeeds?*—is equally important (Figure 1D). In general, it is more challenging to assess potential impact than likelihood of success. But there are two points worth considering.

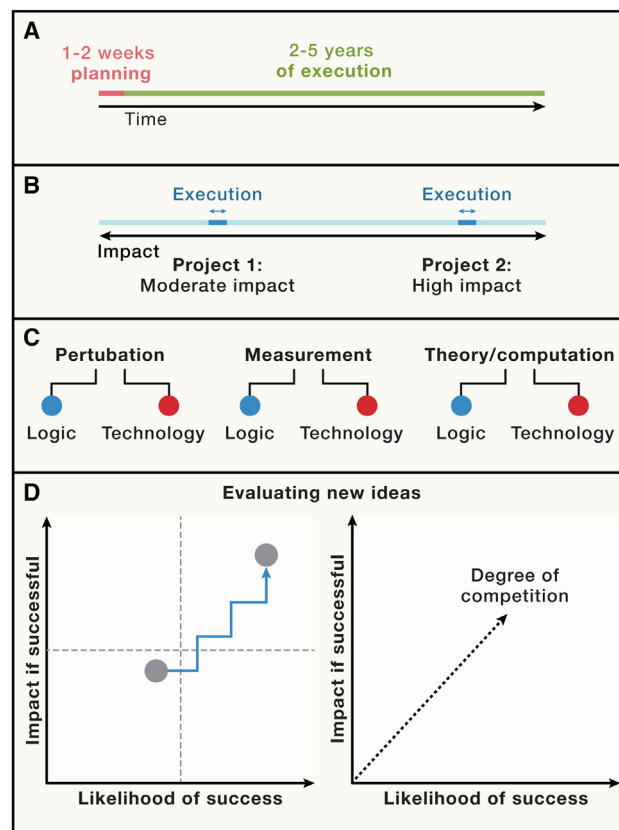


Figure 1. Choosing a good problem

(A) A typical project for an incoming graduate student might involve 1–2 weeks of planning and 2–5 years of execution.

(B) Once you choose a project, you are confined to a relatively narrow band of impact; the solution to a mediocre problem will have incremental impact, whereas solving an important problem will have greater impact. Even if you execute well, it is hard to make the solution to a middling problem interesting.

(C) This is a framework for ideation. A useful oversimplification is that most projects in the biological and chemical sciences involve perturbing a system, measuring it, and then analyzing the data. New ideas generally fall under one of these headings; in each case, they involve an advance in technology or logic. As with any taxonomy, this one is imperfect but useful—you may benefit from learning that your niche is, e.g., perturbation logic or measurement technology.

(D) We find this graph to be a useful framework for evaluating new ideas. The goal is to take an initial idea—the dot in the lower-left quadrant—and improve its likelihood of success and potential impact. You may wish to include a z axis that represents the degree of competition: How many other groups are thinking about or working on this already? What edge do you have? If you aren't the first to solve the problem, how would you pivot? Why don't you go directly to "plan B"?

Table 1. Intuition pumps and common traps

Intuition pump	Examples
<i>Making a one-off perturbation or measurement systematically</i>	Mutating an enzyme and measuring its kinetic parameters Generating a CRISPR mutant and measuring its effect on the transcriptome High-throughput imaging and mass spectrometry
<i>Using the limitations of existing technologies as a muse</i>	Microscopy can't resolve beyond the diffraction limit ... DNA synthesis can't generate a complete genome ... Genetic screens have a precise input (e.g., whole genome knockout set) but an imprecise output (e.g., growth/death) ... Most genetic perturbations consist of knocking out one or two genes; how about combinatorial overexpression (e.g., of transcription factors) ... We can sequence DNA and RNA at high throughput; why not proteins?
<i>I can't imagine a future in which we don't have ____, but it doesn't exist yet</i>	The ability to design highly efficient enzymes in the same way we can currently design other proteins The ability to deliver a genome editing payload to any cell type specifically <i>in vivo</i> The ability to perform 3D tomographic imaging on a live cell
<i>Static → dynamic understanding of biology</i>	Understanding the chain of events between a growth factor binding its receptor and the ensuing cell division event at the same level of detail as turning the key in a car Time-resolved cell atlases and lineage tracing through the entire course of development
<i>Creating a technology platform</i>	Antibody for intracellular targets AI that predicts the set of perturbations required to reach a desired cell state (defined by transcript levels) or phenotype (e.g., glucose-stimulated insulin secretion)
<i>Looking for dogs that don't bark</i>	Why are there no Gram-negative bacteria that live on human skin? Loss-of-function mutations generally render a protein useless; the corresponding gene should be lost over generations. Why are some catalytically inactive enzymes still present and expressed in cells?
<i>Articulating a first step toward addressing a big question</i>	Molecular correlate of consciousness. First step: a genetic screen for genes that, when mutated, produced altered states of consciousness. Printing complete genomes for genome editing and transgenesis. First step: printing and transferring a complete chromosome. Automated synthesis of any molecule of arbitrary structure. First step: automated synthesis of a moderately complex natural product.
Trap to avoid	Description
<i>Recency/familiarity bias</i>	The tendency to choose your next project based on what you just finished or what would be easiest to do, rather than most interesting/impactful.
<i>Being a truffle hound</i>	"Many an academic is like the truffle hound, an animal so trained and bred for one narrow purpose that it is no good at anything else." – Jacob Viner (via Charlie Munger)
<i>Entering a field too late</i>	Most fields/trends follow a sigmoidal curve. A field gets lots of attention at the inflection point of this curve; this is rarely a good time to join.
<i>Self-serving bias</i>	The assumption that when you succeed, it was because of your hard work and intuition, and when you fail, it was due to externalities beyond your control.
<i>Cargo cult fallacy</i>	The mistake of trying to copy the surface-level behaviors of a figure you admire (i.e., working on the same problem), rather than emulating their deeper principles. For a preeminent scientist, the principle to emulate is that they generally became famous by doing something brand new—you should strive to do the same.

First, articulate the criteria by which you hope to be evaluated. These are different for a basic science project than they are for technology development. The distinction might seem obvious, but it is common for trainee vs. principal investigator and author vs. referee conflicts to pivot on a misunderstanding of which category the

project falls under. The root cause is often a failure by one or both parties to articulate this clearly.

For basic science, we suggest the criteria *How much did we learn?* vs. *How general is the object of study?* A high score on either axis is enough to yield substantial impact. For example, the dis-

covery of a ribosome-bound complex that triggers the degradation of stalled polypeptides updates our understanding of translation, and therefore scores well on generality. In complementary fashion, the genomic acrobatics of the single-celled eukaryote *Oxytricha* may not be common to other organisms, but an

elegant effort to map them scores highly on how much we learned ... and may, in turn, yield useful tools for genome editing. The rare projects that score well on both axes tend to be landmark discoveries: two examples are RNA interference and biomolecular condensates.

For technology development, a better alternative is *How widely will it be used?* vs. *How critical is it for the application?* The genomic search tool BLAST is very widely used. Even though it doesn't play a critical role in most projects, its impact has been enormous. A mirror image is cryo-electron tomography, which is too complicated to be used widely but generates stunning data that would be difficult to gather any other way. Technologies that score well on both axes are truly game-changing: not just green fluorescent protein, CRISPR, and AlphaFold but also lentiviral delivery, cell sorting, and massively parallel sequencing, etc.—technologies we cannot imagine living without. Not every project can be the next CRISPR, but the trick is to make sure you score reasonably well on at least one axis. A tool that won't be widely used and isn't critical for an application probably isn't worth building.

In rare cases, it might be appropriate to propose an unusual optimization function. For example, research groups working on frugal science might wish to be assessed by *How many children in low- and middle-income countries now have access to a microscope?* A neglected tropical disease project may be measured by *How many quality-adjusted life years did we save per \$100?*

Second, an assessment of impact doesn't have to be perfect to be useful—something is better than nothing. Even if an estimate of absolute impact is challenging, one can still compare a few options by asking “Which of these would be most impactful if they work?”, and it is far better to make an educated guess than to ignore impact altogether. Moreover, the question of anticipated impact is a useful prompt for a discussion with a senior lab member or the principal investigator. Whether or not your projections align, discuss *why*—how did each of you come to your assessment? The factors you choose reflect your value system and your model of the field, including its future, and are helpful in refining a project idea.

Fix one parameter; let the others float

In generating a new idea, one of the most common failure modes is to fix too many parameters—e.g., the system you will study or the methods you will use. Consider a project that aims to provide a continuous supply of glucagon-like peptide-1 (GLP-1) by engineering a T cell to produce it. This is an interesting idea with some merit, but in our view, too many parameters are fixed. If the most important element of the idea is improving the delivery characteristics of GLP-1 receptor agonists, we should consider any reasonable solution. There are options that are probably more suitable than an engineered T cell; these include peptide engineering to extend half-life or designing an orally available peptide or small molecule. If an engineered cell is meant to be part of the solution, how about a B cell—which can translate and secrete large quantities of protein—instead of a T cell? On the other hand, if the parameter we wish to fix is the use of an engineered T cell, then we should be open to any logical use case—and there are undoubtedly options that are more compelling than GLP-1 production, including the production of smaller quantities of peptides that act locally (e.g., cytokines, chemokines, and growth factors for applications in oncology, autoimmunity, and regenerative medicine).

Which parameter should you fix? This is often determined by a combination of your interests and the expertise of your lab: a trainee in a metabolism lab would fix GLP-1 delivery and be open to alternative methods, while a trainee interested in T cell engineering should fix that parameter and let the payload and application float.

Paradoxically, the opposite can also be true: one can have too few fixed parameters and, as a consequence, too much freedom to think. The statement “I want to do impactful work in cell engineering” is so broad that it can stifle ideation and lead to paralysis. Constraints engender creativity. In class, we point out that next-generation sequencing would have been far simpler if it were an alternative technology that generated fewer, longer reads—i.e., Illumina wasn't the technology we would have asked for, but it's what we got. This constraint has engen-

dered tremendous resourcefulness, not just in computational methods of processing the data (e.g., assembly) but in using sequencing as a readout for gene expression, genomic architecture, and biochemical functions such as protein folding. If you feel stuck, try fixing one parameter at a time and watch your resourcefulness kick into gear.

Learn the “altitude dance”

Projects rarely unfold in a linear fashion; they require frequent course correction. Most trainees should spend more time on a project's decision tree than they currently do. Once you get into a project, you will have learned from your initial experiments, new papers will have been published, and technology will have advanced. As a result, at any decision point, it is rare that you should follow your plan from two years ago; there will likely be a better alternative.

If your genetic efforts to characterize a new phage defense system have hit the skids, instead of troubleshooting them endlessly, why not redo your computational analysis of its distribution in the much larger set of genomes that now exist? Or build an AlphaFold model of each protein and search for other proteins with the same fold? Or print and test a much larger set of candidate systems, given that the cost of DNA synthesis has decreased substantially?

The key to navigating a project's decision tree is to move back and forth frequently between two types of work: getting stuff done (level 1) and evaluating it critically (level 2). These cannot be done at the same time; getting stuff done requires full immersion in the details of experimentation or coding. Critical evaluation demands that you clear your head, step away from the work, and evaluate it as though it were performed by someone else. You can draw concrete conclusions (“What did we learn?”) and then decide what the next step should be, using the tools described in previous sections—e.g., re-evaluating fixed/floating parameters.

Often, you need to troubleshoot what went wrong so you can get back to the original plan; this is reasonable up to a point. But too much of this, coupled to a failure to seek alternative solutions, is the essence of being stuck in a rut. Nor

Box 1. Assumption analysis. For a project in question, list every assumption you would make from its inception through its conclusion. Assign two scores to each one: likelihood of success and duration of effort. After creating a draft, we encourage the trainee to review it with a senior graduate student or postdoc to identify missing assumptions and adjust risk and timing scores. This is a great way to learn quickly, and the document facilitates concrete feedback that is more useful than simply asking, “What do you think of my idea?” When the assumption analysis has been refined, a trainee can share it with their principal investigator. This will help align expectations around where the project is headed and how much risk each step entails. If either party is uncomfortable with the plan, it will be easier to understand why and adjust accordingly (or decide to vet a different idea). We note that there are two types of assumptions: those about the underlying reality (e.g., the existence of a cell type) and those about the capabilities of technology (e.g., that one can isolate them). The latter might change during the course of a project while the former will not. *1 = low, 5 = high.

Assumption	Risk*	Time (mo)
There exist new enteroendocrine cell types beyond what is already known (L and M cells).	1	6
Epithelial cells can be isolated from subepithelial intestinal cells.	1	8
Epithelial cells can be dissociated into a single-cell suspension.	1	8
scRNAseq will generate high-quality transcriptomic data from dissociated epithelial cell suspensions.	2	12
scRNAseq data analysis (e.g., by dimensionality reduction) will be sufficient to categorize known intestinal epithelial cells by type, including L and M cells.	2	16
Such an analysis of scRNAseq data from the murine intestine will reveal new cell types.	3	20
For several of these cell types, it will be possible to predict a set of markers and a function that can be tested experimentally.	4	24
The existence of 2–3 new cell types can be validated experimentally.	5	24
At least one of these cell types can be selectively knocked out of a mouse.	3	26
The knockout will yield a biologically relevant phenotype.	5	30

does success relieve you of your obligation to engage the decision tree; small victories can be equally opportune moments to pivot to a more interesting or speedier plan.

Some people have a predilection for gathering data or writing code, but rarely stop to consider its implications and upgrade their project plan. Others are brilliant strategists but have trouble rolling up their sleeves and reducing plans to practice. In our experience, the most successful scientists move back and forth frequently between planning and doing.

Capitalize on the “adversity feature”

Some mountain biking trails have a field of boulders to traverse or a narrow wooden bridge to cross. You and I might think of these as obstacles, but seasoned cyclists call them the “rock garden” and the “log ride.” They think of the trail as a game and the element of adversity as an opportunity to develop a skill.

Adversity in a project should be viewed in the same way: inevitable and opportune. The inevitability cannot be overstated: almost every project suffers an existential crisis or takes a sharp turn. Odds are that yours will, too. It will take you by surprise, but it shouldn't!

Happily, there is a silver lining. A crisis in your project is an opportunity to do two things of great value: 1) Fix the problem and upgrade your project at the same time—make it better than it was before. 2) Realize you are backed into a corner and reason your way out—one of the best growth opportunities on the training menu. This can be hard to remember after the pain of a failed experiment or a paper that scoops you. But this is the crisis you have been waiting for: do not waste it.

One place to start is to acknowledge, at the outset of a project, that a singularly mapped path has virtually no chance of coming true. You should understand that you are picking an *ensemble* of possible projects that have a broad chance of yielding impactful results. When (inevitably) you hit a roadblock, you will need to be flexible about your fixed vs. floating parameters and optimization function and will likely end up on another path within this ensemble.

Turn a problem on its head

There are several ways to navigate around a problem. Three are particularly notable; the first goes back to the idea of fixed parameters. During the ideation process, we encouraged against fixing too many parameters to avoid a poor technique-appli-

cation match. But as a project launches and gains momentum, additional parameters naturally get fixed. For example, you might be 1) using spatial transcriptomics to 2) study the interactions between antigen-presenting cells and T cells 3) in the tumor microenvironment. This is not inherently a problem—without fixing parameters, you can't run an experiment! But when the time comes to troubleshoot, a useful way to start is to make a list of the fixed parameters and then let each of them float, one at a time, to explore alternative paths around a roadblock. This is especially important when you are in a deep rut; the solution is often to let a “sacred” fixed parameter float.

The second strategy is to turn a problem on its head. In an illustrative example,³ the initial idea was to develop small-molecule degraders of two kinases by conjugating known ligands for each protein to pomalidomide, which recruits the ubiquitin ligase cereblon. However, it simply did not work. Instead of trying to force a solution, the authors identified the core problem: that any individual kinase may resist chemically induced degradation. They turned it on its head by asking instead “Which kinases *can* be degraded?” To answer this question, they created a promiscuous small-molecule kinase degrader and used it to find 28 kinases that were degradable,

including targets that are biologically relevant and therapeutically important. Although they didn't accomplish their original goal, their new project demonstrated an important and general learning that target engagement is necessary but not sufficient for degradation, defined a large set of kinases that are druggable by degradation, and arguably had more impact than if their original plan had succeeded.

A way to extend this strategy can be used when you carry out a project, find the answer, and realize it no longer applies to the question you started with ("I have the answer; what is the question?"). Consider a project that aims to identify the receptor for a new steroid hormone. The data reveal that the molecule binds multiple nuclear hormone receptors with varying affinities. The original question cannot be answered; there is no single receptor. So, what *is* the question to which the data provide an answer? One option: how does a finite pool of nuclear hormone receptors sense a near-infinite set of lipids and steroids? The data might suggest that the answer involves combinatorial sensing: each molecule is bound by multiple receptors at different affinities,

in a way that forms a unique pattern like a piano chord.

Conclusion

There is no right way to choose a problem, but we hope this piece provides a starting point—organizing principles for a more systematic way to go about it. Nothing would make us happier than to spark a discussion and perhaps inspire others to teach the same course. We will happily offer course materials to anyone who is interested.

ACKNOWLEDGMENTS

I developed and taught this course with my former PhD advisor, Christopher T. Walsh, who passed away just as we began writing this piece. Nearly every idea described above was the product of robust discussion between us. As such, I consider him a co-author on this piece and owe him an immeasurable debt of gratitude. We are deeply indebted to Bryan Roberts, Ruslan Medzhitov, John Stuelpnagel, and Elliot Hershberg for ideas that have become central tenets of the course and to Julia Bauman, Xiaojing Gao, Zemer Gitai, Elliot Hershberg, Wendell Lim, Ruslan Medzhitov, and Wesley Sundquist for constructive feedback and new ideas that were incorporated prior to submission. We are grateful to the many students who

have taken BIOE 395 and to our colleagues who participated in the first PI-focused round of the course (Polly Fordyce, Lacia Bintu, Elizabeth Sattely, Hawa Racine Thiam, Rogelio Hernandez-Lopez, Jennifer Brophy, Steven Banik, Nicole Martinez, Florentine Rutaganira, and Erin Chen); their feedback has been immeasurably helpful in shaping our thinking and improving the course.

DECLARATION OF INTERESTS

M.A.F. is a co-founder of Kelonia and Revolution Medicines, a member of the scientific advisory boards of the Chan Zuckerberg Initiative, Stand Up to Cancer, NGM Biopharmaceuticals, and TCG Labs/Soleil Labs, and an innovation partner at The Column Group.

REFERENCES

1. Alon, U. (2009). How to choose a good scientific problem. *Mol. Cell.* 35, 726–728. <https://doi.org/10.1016/j.molcel.2009.09.013>.
2. Dennett, D.C. (2013). *Intuition Pumps and Other Tools for Thinking* (W. W. Norton & Company).
3. Huang, H.-T., Dobrovolsky, D., Paulk, J., Yang, G., Weisberg, E.L., Doctor, Z.M., Buckley, D.L., Cho, J.-H., Ko, E., Jang, J., et al. (2018). A Chemoproteomic Approach to Query the Degradable Kinome Using a Multi-kinase Degradator. *Cell Chem. Biol.* 25, 88–99.e6. e6. <https://doi.org/10.1016/j.chembiol.2017.10.005>.