Welcome to:

Protein Engineering and Design

Parisa Hosseinzadeh

Winter 2022

Introduction

Name
Pronouns
Lab/year
What are you hoping to learn from this class
How do you evaluate your coding/biochemistry skills

What are you looking forward to the most in 2022

Class core values

1. Be **respect**ful to yourself and others
2. Be **confident** and believe in yourself
3. Always do your **best**
4. Be **cooperative**
5. Be **creative**
6. Have **fun**
7. Be **patient** with yourself while you learn
8. Don't be shy to **ask "stupid" questions**

About our class

What to bring?
- Yourself
- Your attention and curiosity
- Your laptops

About our class

What to bring?
    Yourself
    Your attention and curiosity
    Your laptops

What do I need to know?
    What is a protein
    Protein secondary structure
    What is an amino acid
    General of protein folding

About our class

Pre-class assessments (So you think you know X)

     Quiz to check your knowledge on the topic
     Short videos/readings that will bring you up to speed

About our class

Pre-class assessments (So you think you know X)
    Quiz to check your knowledge on the topic
    Short videos/readings that will bring you up to speed

Post-class assignments
    Short quizzes
    Assay questions
    Readings (for journal club)
    Running a guided code

    …

Activities

Pre and post class assignments
In class activities
      Asking questions, submitting in-class work, presence
      Either through canvas or google colab (gmail account)
      1 session allowed absence (excluding first session, ask me if you need more)

Activities

Pre and post class assignments

In class activities

Journal presentation

    Moderators – give backup information, take care of organization, ask questions

    Everyone else – present figures

Activities

Pre and post class assignments
In class activities
Journal presentation
Proposal write-up
>   1-page describing your aims and ideas (410)
>   1-page describing your aims and ideas + 1-2 page method description (510)

Activities

Pre and post class assignments
In class activities
Journal presentation
Proposal write-up
Mock study section
    Grant review and feedback

Panels/Guests

Expert lectures on protein structure prediction
      Dr. Afua Nyarko (OSU)
      Dr. Richard Cooley (OSU)
      Dr. Elizabeth Kellog (Cornell)

Panels/Guests

Expert lectures on protein structure prediction
      Dr. Afua Nyarko (OSU)
      Dr. Richard Cooley (OSU)
      Dr. Elizabeth Kellog (Cornell)
Industry panel
      Arzeda, Novo Nordisk, Bayers, Ginkgo Bioworks

Panels/Guests

Expert lectures on protein structure prediction
  Dr. Afua Nyarko (OSU)
  Dr. Richard Cooley (OSU)
  Dr. Elizabeth Kellog (Cornell)
Industry panel
  Arzeda, Novo Nordisk, Bayers, Ginkgo Bioworks
Startup CEO
  Dr. Daniel Adriano Silva
  (previous CEO of Neoleukin, now CEO at Monodbio)

Feedback

Assignments and activities

      Detailed feedback provided by instructor/TAs
      Answering questions during hands-on sessions
      Help navigate the paper for moderating group
      Office hours: Tuesdays 9-10 AM

Feedback

Assignments and activities

      Detailed feedback provided by instructor

      Answering questions during hands-on sessions

      Help navigate the paper for moderating group

      Office hours: Tuesdays 9-10 AM

Feedback from you

      Mid-term and final feedback

      Personal feedback (email, anonymous feedback)

      2 + 1 feedback every two weeks

Feedback

Assignments and activities
      Detailed feedback provided by instructor
      Answering questions during hands-on sessions
      Help navigate the paper for moderating group
      Office hours: Tuesdays 9-10 AM

Feedback from you
      Mid-term and final feedback
      Personal feedback (email, anonymous feedback)
      2 + 1 feedback every two weeks

Slack ?

Questions?

Week 1, Lecture 1

Protein engineering:
The why, what and how

Learning Objectives

1. Describe goals of protein engineering and its applications
2. Identify proper method for different applications of protein engineering methods
3. Critically evaluate the application of protein engineering in different scenarios

Protein Engineering

What is protein engineering?

Protein Engineering

What is protein engineering?

Why do we need to engineer proteins?

Engineering proteins is an old concept

> 20,000

3 articles:
studying the effect of N-substituted amino acids

1949

2020
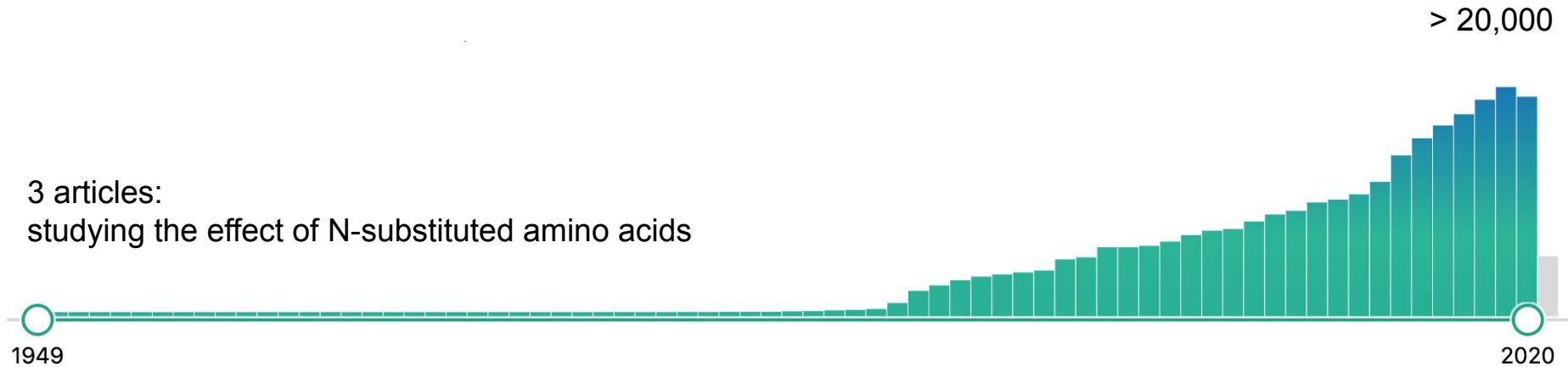
Early engineering attempts was focused on understanding the role of natural variants

1949   1972
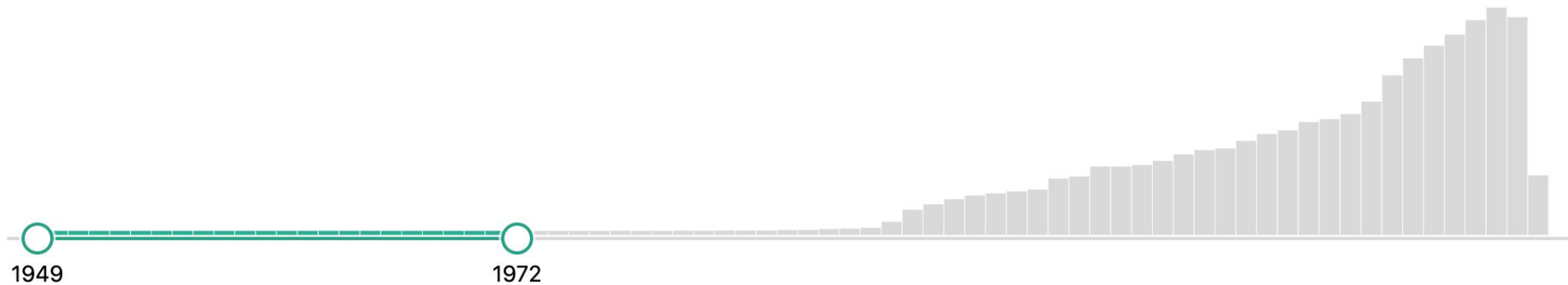
Rise of recombinant DNA technology was a hallmark in protein engineering

Recombinant DNA technology

1949   1972
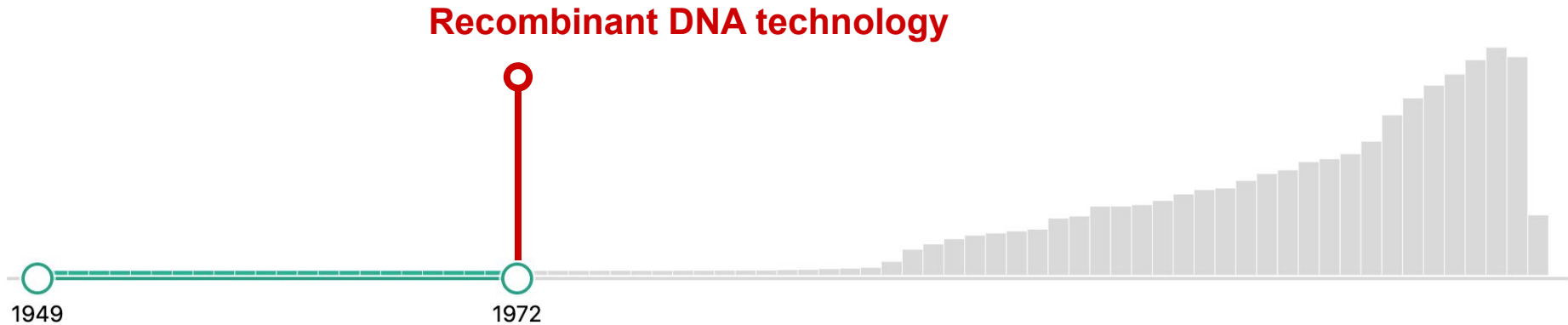
Recombinant protein expression allowed researchers to express proteins in high yields

1972   1982
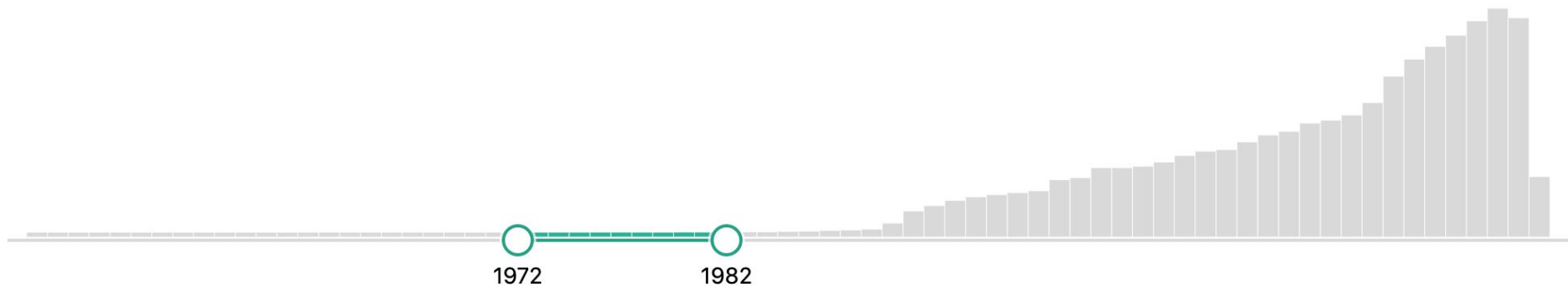
The idea of designing new proteins emerged with the ability to express proteins

Whither protein structures?

from S. D. Dover

THE pipe-dream of designing and producing a protein to carry out a specific task has been brought nearer reality by the advent of recombinant DNA which offers the possibility of isolating, mutating and cloning particular genes. The resulting change in amino acid sequence specified could produce the intended structural change in the protein. But this will only be possible with an adequate theory of the effect of sequence changes on structure, based firmly on experimental structural information. Without this
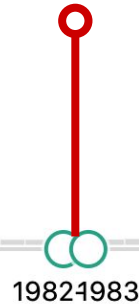
1982 1983

The idea of designing new proteins emerged with the ability to express proteins

Abstract

The prospects for protein engineering, including the roles of x-ray crystallography, chemical synthesis of DNA, and computer modelling of protein structure and folding, are discussed. It is now possible to attempt to modify many different properties of proteins by combining information on crystal structure and protein chemistry with artificial gene synthesis. Such techniques offer the potential for altering protein structure and function in ways not possible by any other method.
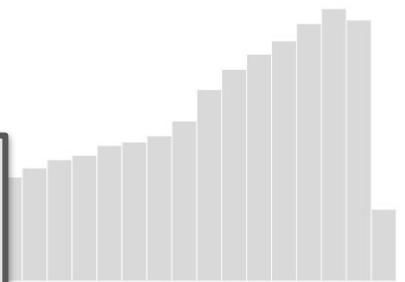
The idea of designing new proteins emerged with the ability to express proteins
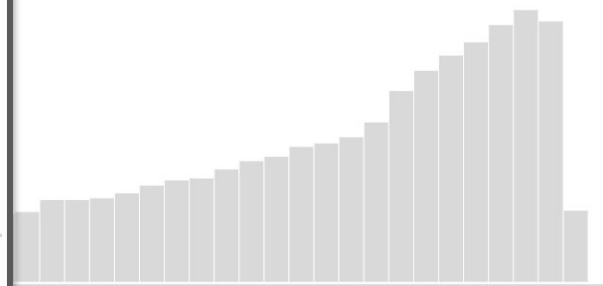
Enzyme engineering

William H. Rastetter ✉

Abstract

Enzyme research and development efforts have been shaped by the tools and concepts available for enzyme production and utilization. A new phase of enzymology characterized by the production of modified protein catalysts has begun, made possible by recombinant DNA technology.

Four methods of protein engineering

Rational Design
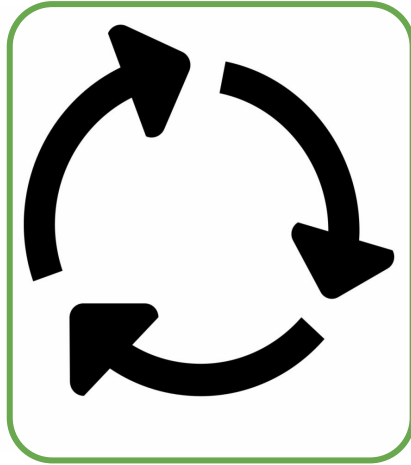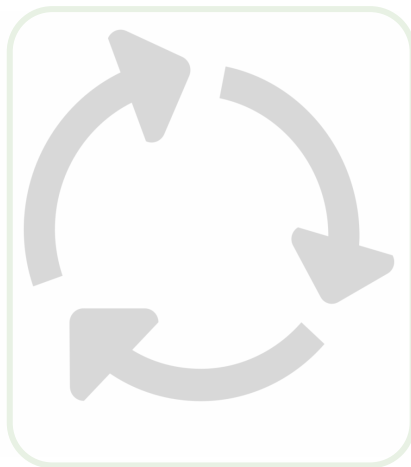
Directed Evolution

De Novo Design

Machine Learning

Rational Design

uses prior knowledge, often structure guided to engineer proteins

Rational Design

Directed Evolution
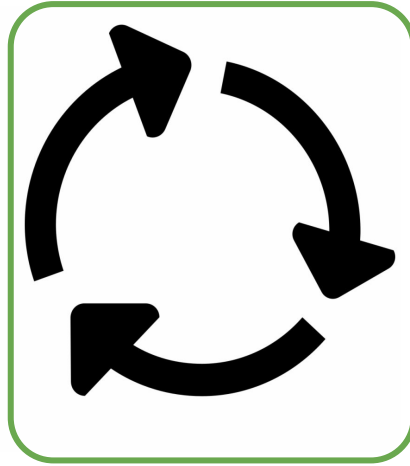
De Novo Design

Machine Learning

Directed Evolution

uses a process of mutation and selection, inspired by evolution.

Rational Design   Directed Evolution   De Novo Design   Machine Learning
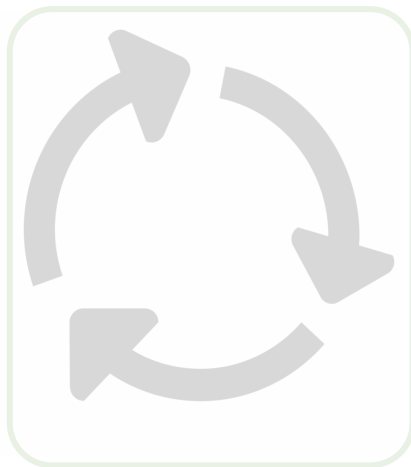
de Novo Design

creates <u>new</u> proteins often by physics-based and heuristic methods

Rational Design

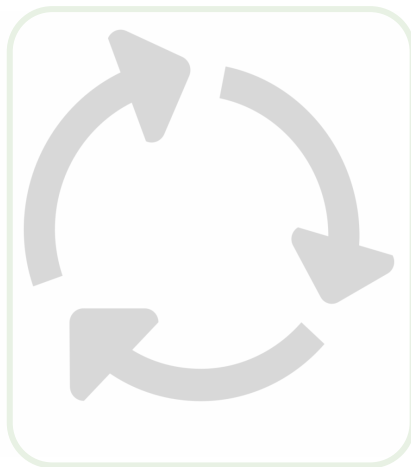Directed Evolution

De Novo Design

Machine Learning

Machine Learning
learns from large databases for prediction, classification, generation

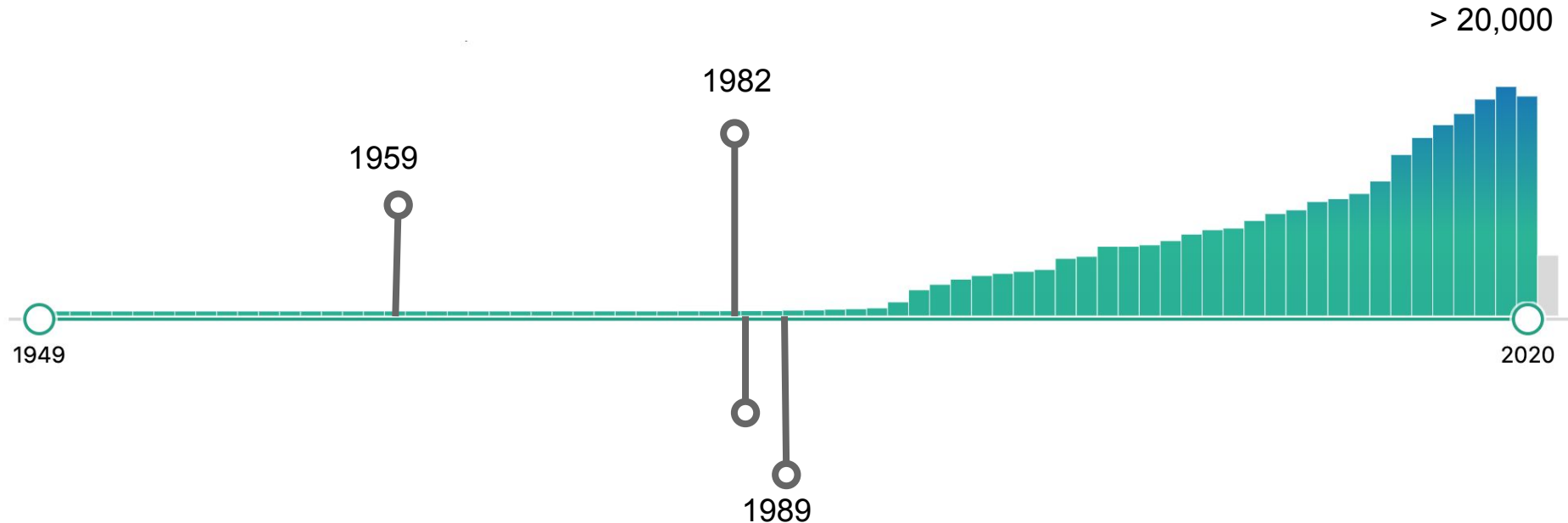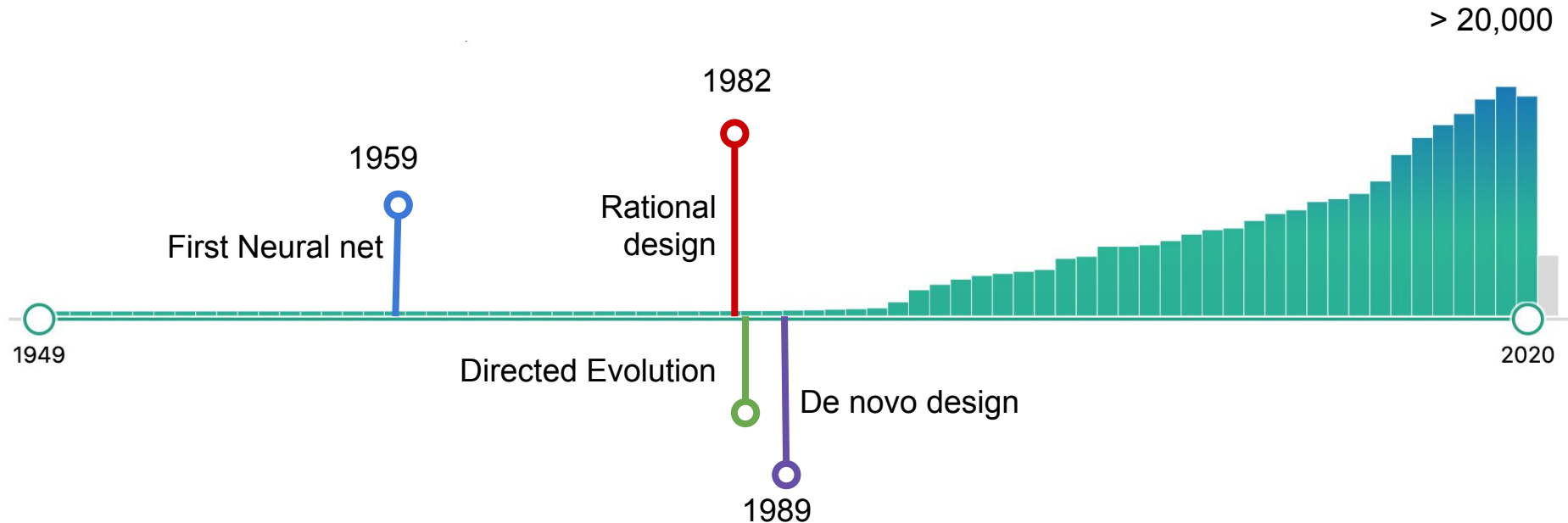Rational Design

Directed Evolution

De Novo Design

Machine Learning

Four methods of protein engineering

> 20,000

1982

1959

1949   2020

1989

Four methods of protein engineering

> 20,000

1982

1959

Rational design

First Neural net

Directed Evolution

De novo design

1949

2020

1989

# Each method has its own advantages and disadvantages

uses prior knowledge, often structure guided to engineer proteins

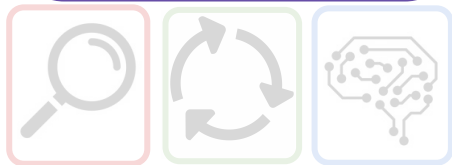| Advantage	Disadvantage
We're building off of what works → higher chance of success	Limited to what we know
Simpler if you have a good guess	Limited by what exists in nature
Requires fewer resources	You need to know the structure

Each method has its own advantages and disadvantages
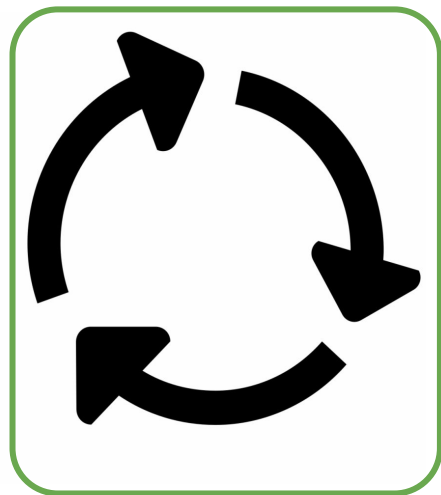
uses a process of mutation and selection, inspired by evolution

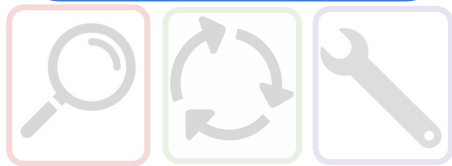| Advantage	Disadvantage
High throughput and can test many things at once	Sequence space is very large and not everything can be sampled
Don't need prior knowledge	Is not interpretable (we don't know why things worked)
Can screen multiple features at once	We're not sure if we're getting the best result
Random mutations = casting a wider net to make sure things work	You need to have a selection/screening method and you have to have some minimal activity

Each method has its own advantages and disadvantages

creates <u>new</u> proteins often by physics-based and heuristic methods

Advantage	Disadvantage
If it works, it generates really good proteins because it considers all the rules	Low success rate
Novelty = less chance of resistance/defense	novel structures → may cause unknown problems downstream
Can create novel function/structure and is not limited to what we know	While not dependent on prior knowledge, still needs the rules
Great for specificity and full control	

Each method has its own advantages and disadvantages

learns from large databases for prediction, classification, generation

Advantage	Disadvantage
The computational power is getting better	Needs large amount of data
It can uncover new patterns and rules that we can't	Can't be specific/target-based (great for general cases)
Fully reproducible	Can sample new sequences but in the neighborhood of what exist
Can take in multiple considerations into account	

Match methods to applications

Developing a biological logic gate
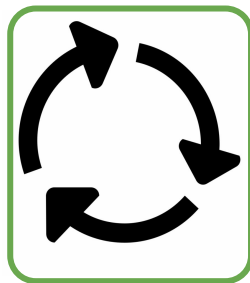
Generating a protein that binds to another protein

Stabilizing an enzyme with known structure

Generating a protein with a novel fold

Enhancing the activity of an enzyme

Generating a fluorescent protein smaller than GFP

For the next lecture:

1. Pre-class assessment for the next lecture
        Needs to be done before the start of class, will be available after this class

2. Post-class assignment – Due next week

3. Make sure you have PyMol installed for the next lecture

Next lecture:

It's all about structure!