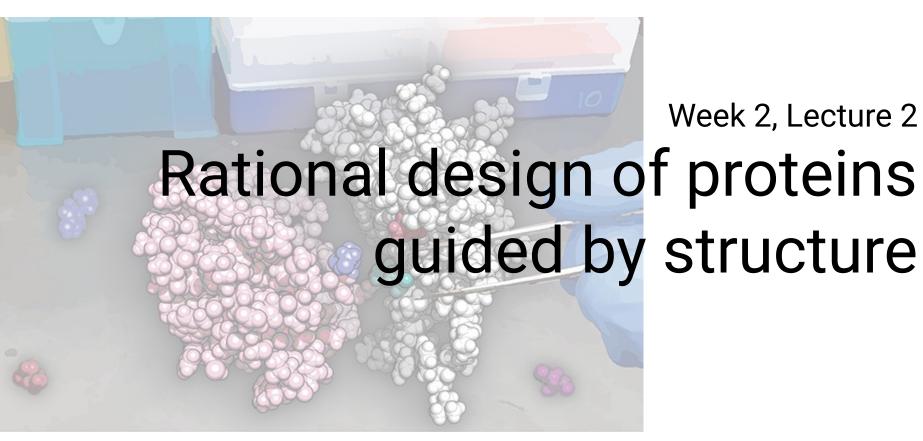
#### Class core values

- 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**





Week 2, Lecture 2

#### Learning Objectives

- Describe the concept of rational design
- 2. Identify potential mutations to increase stability
- 3. Describe different methods for rational design
- Apply tools for structure analysis, evolutionary information, and computational techniques to design
- 5. Identify the limitations of rational design approaches



#### Rational design





#### Rational design is often guided by prior knowledge, often structure











## Rational design is often guided by prior knowledge, often structure



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



# The next step is to consider the features we want to improve (the what)



# Enhancing proteins stability is one of the most commonly used engineering applications

Structure-guided mutagenesis based on prior knowledge to improve features



Thermal stability

pH stability

Oxidant stability



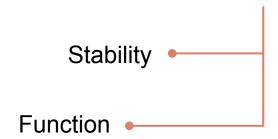
# Another commonly used application of rational design is enhancement of function/activity

Structure-guided mutagenesis based on prior knowledge to improve features

Improved activity

Substrate specificity

Binding selectivity





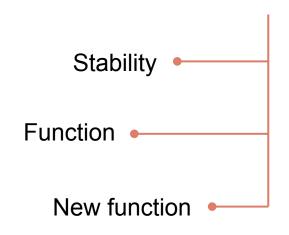
#### Rational design can also be used to design biomimetics or create new function

Structure-guided mutagenesis based on prior knowledge to improve features

Mimicking other proteins

Changing substrates

Changing cofactors





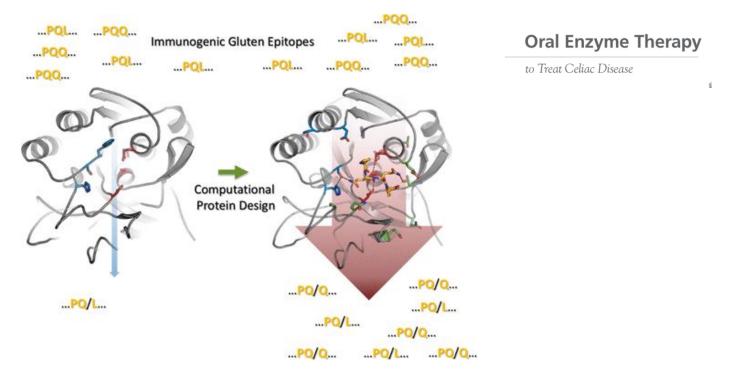


**Oral Enzyme Therapy** 

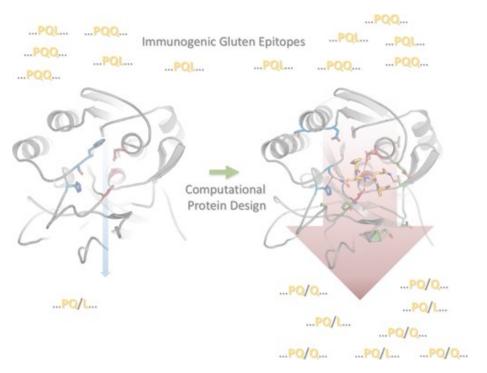
to Treat Celiac Disease

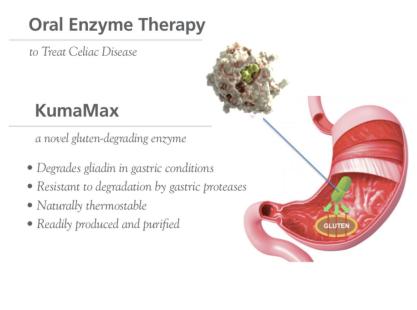






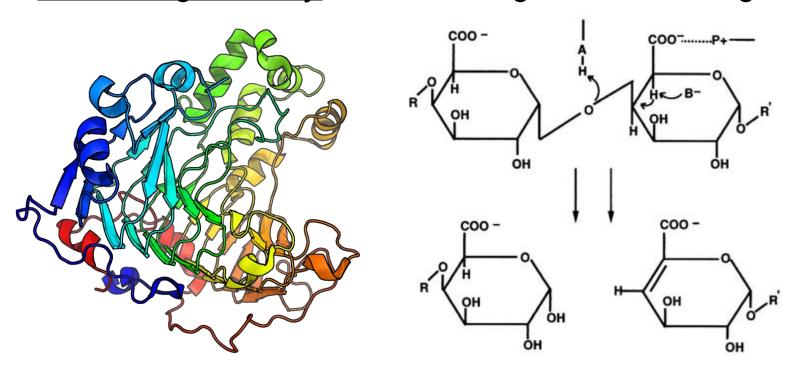






#### In-class activity:

#### Enhancing stability of PelN using Rational Design



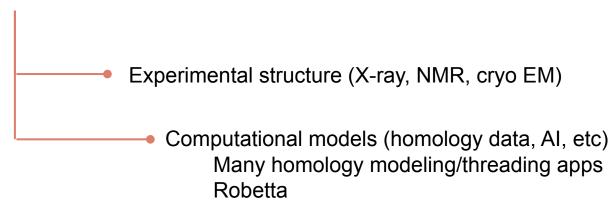


## Many databases are available that include protein structures with statistics about quality

```
Experimental structure (X-ray, NMR, cryo EM)
PDB
CSD
EMBL-EBI
```

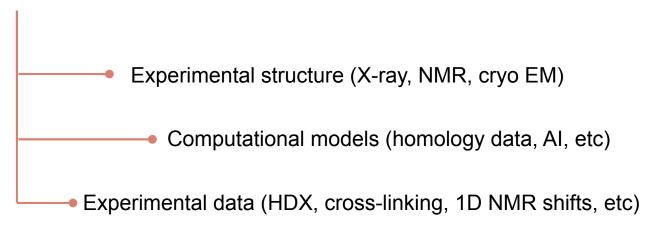


## Computational modeling methods offer great insight about proteins without validated structure





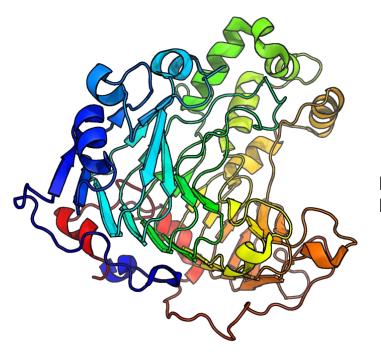
# For more challenging cases, experimental data can provide a powerful intuition





#### In-class activity:

#### Enhancing stability of PelN using Rational Design



PelN

PDB: 5GT5

# The engineering process is often guided by some prior knowledge



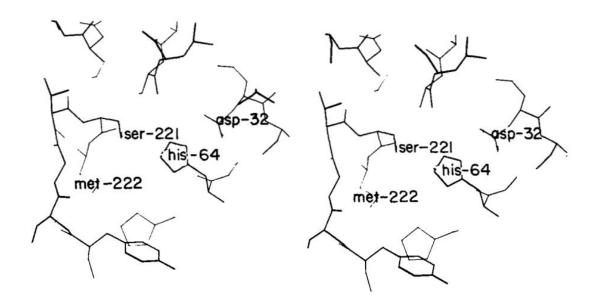
## Our biophysical/biochemical knowledge can guide us in finding beneficial mutations

Structure-guided mutagenesis based on **prior knowledge** to improve features

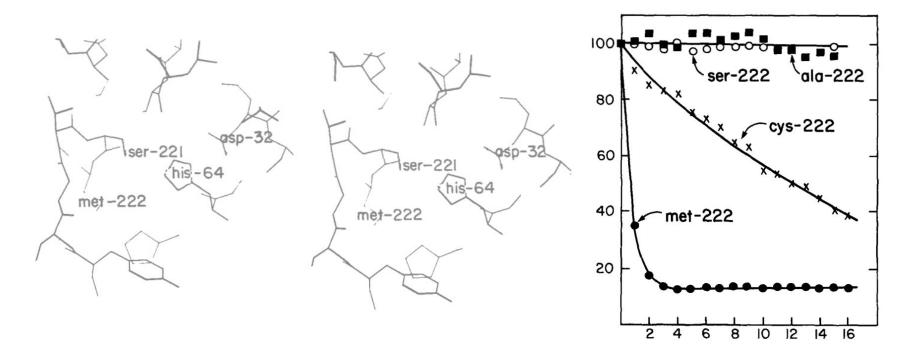
Biophysical and Biochemical knowledge (aka physics-based)



## Observations about amino acid properties can lead to significant outcome: the case of <u>Subtilisin</u>



## Observations about amino acid properties can lead to significant outcome: the case of <u>Subtilisin</u>





## Our biophysical/biochemical knowledge can guide us in finding beneficial mutations

Structure-guided mutagenesis based on **prior knowledge** to improve features

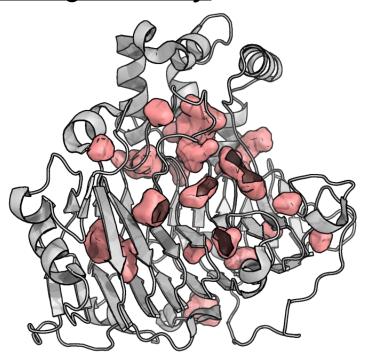
Biophysical and Biochemical knowledge (aka physics-based)

- Burying polar residues in the core of a protein has a huge energy cost
- New hydrogen bonds are energetically favorable
- The size of an enzyme pocket affects its substrate scope
- ...



#### In-class activity:

#### Enhancing stability of PelN using Rational Design



Cavities in PelN PDB: 5GT5



## Evolutionary data and close relatives contain rich information about protein fold/function

```
Biophysical and Biochemical knowledge

Statistical and evolutionary information

(aka heuristic)
```



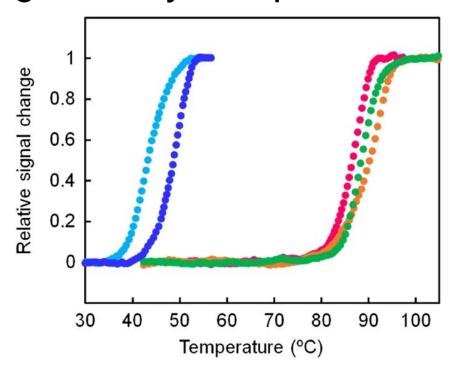
## Ancestral sequence reconstitution can be used for enhancing stability of a protein

anc IPMDH-IQ M-MTYKI AVLPGDGI GPEVVAEAVKVLEAVAE-KYGLEFEFEEAL VGGAAI DATGTPLPE anc IPMDH-ML M-MTYKI AVLPGDGI GPEVVAEAVKVLEAVAE-KFGLEFEFEEAL IGGAAI DATGTPLPE T. thermophilus IPMDH M---KVAVLPGDGI GPEVTEAALKVLRALDE-AEGLGLAYEVFPFGGAAI DAFGEPFPE B. subtilis IPMDH M-KKRI ALLPGDGI GPEVLESATDVLKSVAE-RFNHEFEFEYGL IGGAAI DEHHNPLPE S. cerevisiae IPMDH MSAPKKI VVLPGDHVGQEI TAEAIKVLKAI SDVRSNVKFDFENHL IGGAAI DATGVPLPD

Part of an MSA (Multiple Sequence Alignment)



### Ancestral sequence reconstitution can be used for enhancing stability of a protein





## Evolutionary data and close relatives contain rich information about protein fold/function

```
Biophysical and Biochemical knowledge

Statistical and evolutionary information

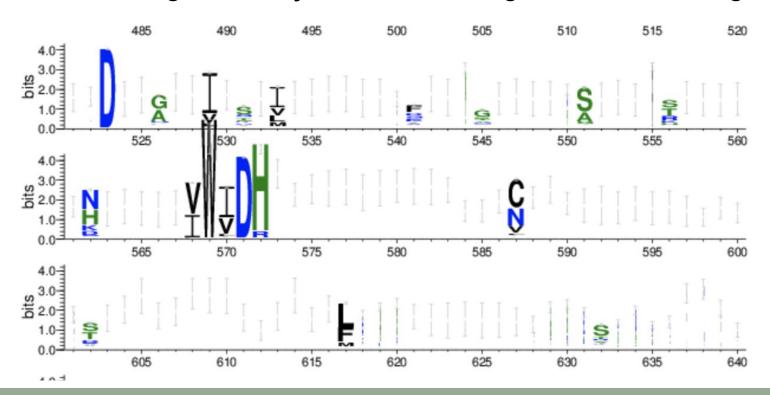
(aka heuristic)
```

- Multiple Sequence Alignment
  - o NCBI, JackHmmer, Pfam



#### In-class activity:

#### Enhancing stability of PelN using Rational Design





## Structural features from one protein can be grafted onto other proteins to exert new function

Structure-guided mutagenesis based on **prior knowledge** to improve features

Biophysical and Biochemical knowledge
(aka physics-based)

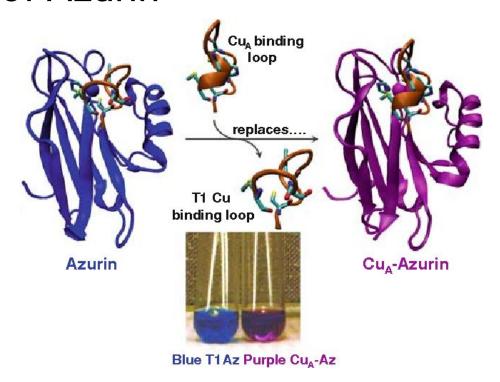
Statistical and evolutionary information
(aka heuristic)

Grafting known motifs

- Grafting loops onto structures
- Adding binding ligands inside other cavities
- Placing binding motifs onto other scaffolds
- ...



#### Loop grafting was used to change metal binding features of Azurin





## Visual or computational screening can help narrow down the number of experiments

Structure-guided mutagenesis based on prior knowledge to **improve** features

Virtual screening •

Visual observations Computational Methods



#### In-class activity:

Enhancing stability of PelN using Rational Design





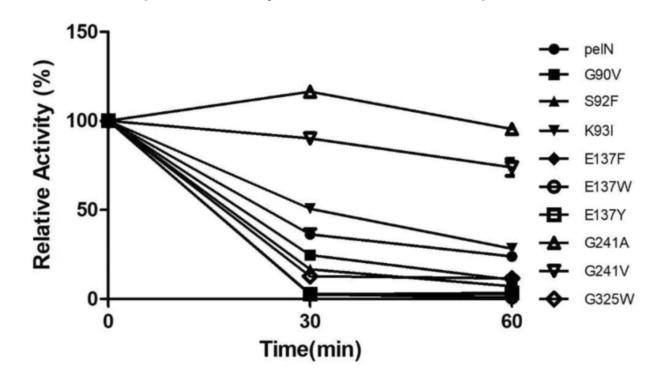
## Experimental testing is the ultimate method for assessing the effectiveness of our designs

```
Virtual screening

Experimental testing
```



#### In-class activity: <u>Enhancing stability</u> of PelN using Rational Design





#### Possible search space is extremely large

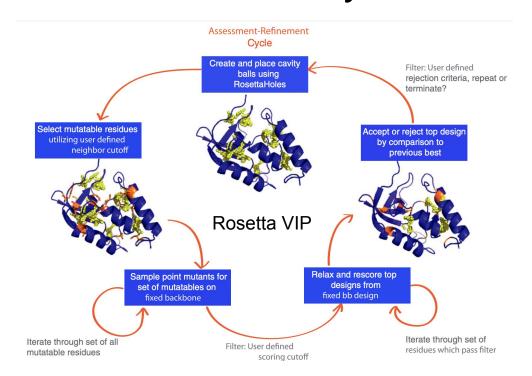


Possible search space is extremely large =>

Computational methods can be used to speed up the rational design process

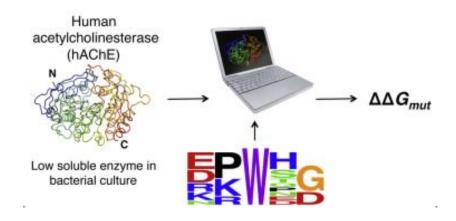


### Automated search can be used to fill in core cavities and enhance stability



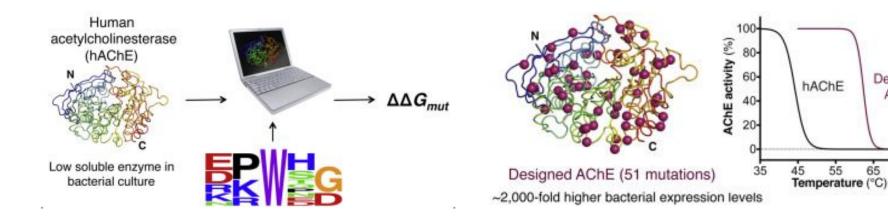


### MSA information can be automatically used to find stabilizing mutations





## MSA information can be automatically used to find stabilizing mutations

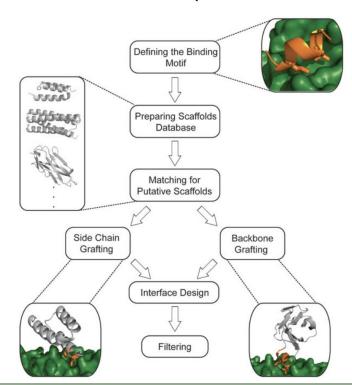




Designed

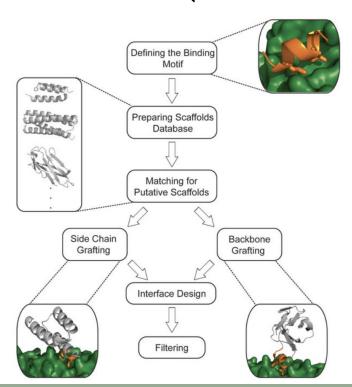
AChE

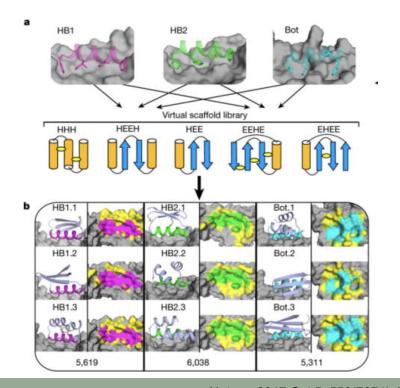
### Secondary structure motif grafting can be fully automated (Rosetta motif graft)





### Secondary structure motif grafting can be fully automated (Rosetta motif graft)

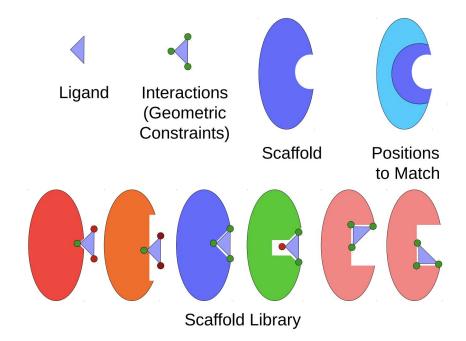






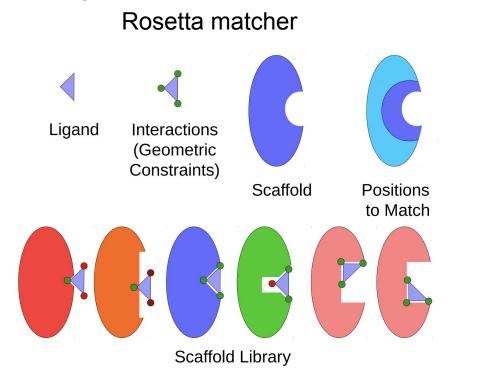
#### Interacting residues required for binding can also be placed in new scaffolds in an automated way

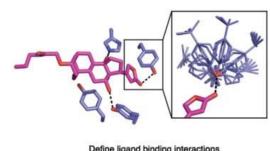
Rosetta matcher



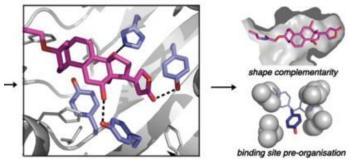


#### Interacting residues required for binding can also be placed in new scaffolds in an automated way





Define ligand binding interactions



place ligand and interacting residues in scaffolds & design binding site sequence

select pre-organised sites with high shape complementarity



#### For the next lecture:

- 1. Proposal draft: Due next week (discussion in class)
- 2. First journal: Will be discussed next week
- 3. Post-class assignment for this lecture: Rosetta-guided SSM
  - a. Due next week



### Next lecture: A designed heme-[4FE-4S] metalloenzyme

