Best practices in online teaching for BMB classrooms

Brought to you by
July 10 | 12–4:30 p.m. EDT

Virtual lab instruction: Virtual CUREs in a biochemistry laboratory

1.00-1.45pm

Ellis Bell, University of San Diego,

Anthony Bell, University of San Diego

Betsy Martinez-Vaz, Hamline University

Tamara Mans, North Hennepin Community College

Many High Impact Educational Practices tend to involve:

"learning at the edge of chaos"

(Bertschinger and Natschläger 2004, Kleiman 2011)

photo courtesy of Rick Parent

Road Map:

Brief Background to CUREs

Overview of a Full Semester Virtual CURE

Overview of a Virtual Modular CURE within the Semester

Computational Approaches to Incorporate into your CURE
What Questions do you want to ask?
How to Organize the lab periods

Integrating with available wet lab approaches and data

Working as Part of a bigger project: eg MCC Malate Dehydrogenase projects

Building in Collaboration: **Drs Betsy Martinez-Vaz & Tamara Mans**

Adapting to any protein you want to use: **Dr Anthony Bell**

This course has five specific learning goals:

i)   *Find, use and present* relevant primary literature, protein sequences, structures and bioinformatics tools

ii) Understand the various roles that non- covalent interactions may play in the structure, function and experimental analysis of an enzyme.

iii) Keep an *accurate laboratory notebook* that allows others to interpret and reproduce reported experiments, and work as an effective team.

iv) Be able to *develop a hypothesis* and *research proposal* and design and perform experiments to interrogate your hypothesis, &

v) Can present the basis and results of your hypothesis/project using verbal, visual or written media to a variety of audiences, and *draw evidence based conclusions* using data obtained from *a variety of biochemical and biophysical techniques* that explore protein structure-function relationships.
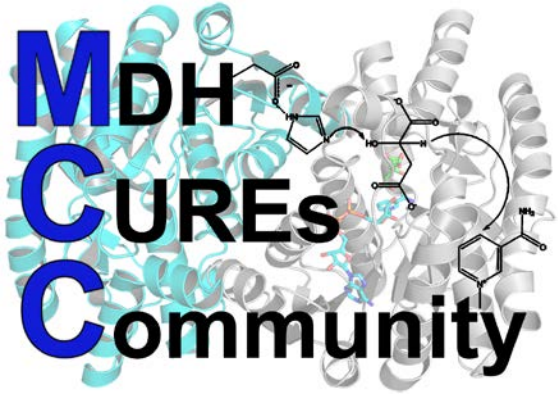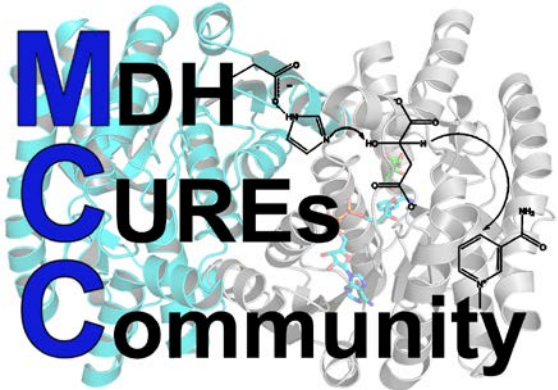
Part 1: Overview: What is a CURE?

Key Elements of a CURE

Learn/Practice Techniques, Reproduce control (WT) data, Learn Data Analysis approaches
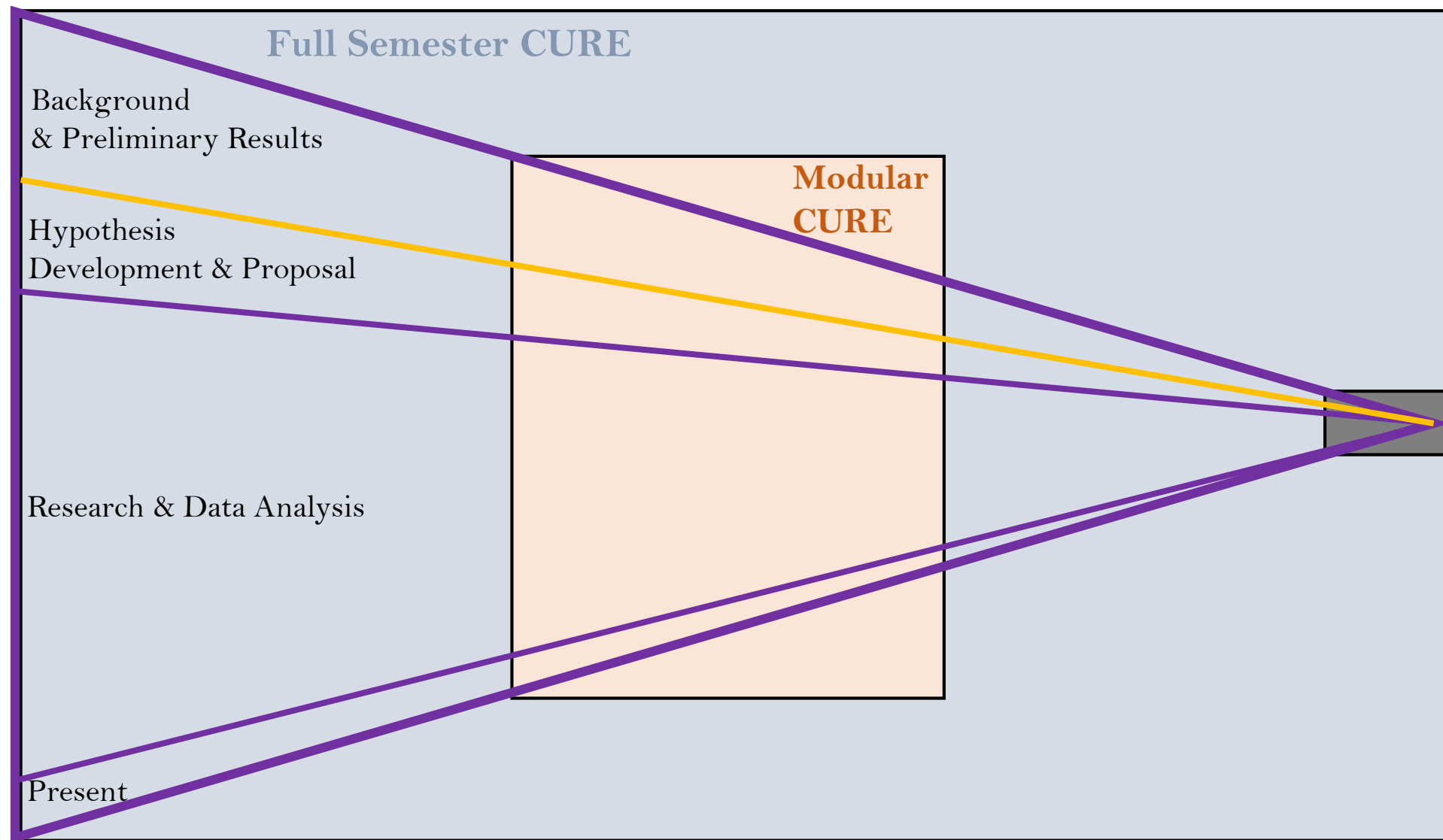
{ Relevance
Scientific Background }
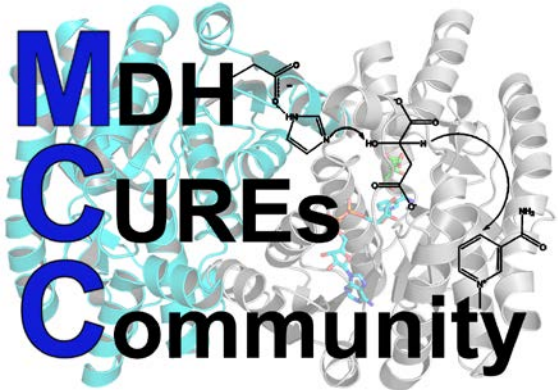Hypothesis Development
Proposal }

Experiments, Teamwork, Collaboration, Reproducibility
Data Analysis & Drawing **Evidence Based Conclusion** }

Presentation: Written, Oral, Poster etc

Overview: How long is a CURE?

Full Semester CURE

Modular CURE

Background & Preliminary Results

Hypothesis Development & Proposal

Research & Data Analysis

Present

Potential Schedule for 14 week semester- full semester CURE

Semester Week #	Activity	Student Assessment
1	Introduction to the semester and project etc, Notebooks, Lab Safety Training	
2	Background: finding and reading the literature	Proposal Introduction
3	Bioinformatics: Clustal etc	Draft Hypothesis
4	Molecular Visualization & Hypothesis Development	Refine hypothesis, draft proposal
5	Protein Preparation- Mutant Design, make proteins	Proposal Presentation
6	Characterize Proteins	Journal Club
7	Project	Write up Methods
8	Project	Journal Club
9	Project	Write up Methods
10	Data Analysis and Presentation	Figures and Tables etc
11	Project continued	Journal Club
12	Project continued	Write up Methods
13	Project continued: conclusions etc	Draft final report/presentation
14	Final Presentation	

Potential Schedule for 6 week Modular CURE in the Semester

Semester Week #	Activity	Student Assessment
1	Introduction to the semester and project etc, Notebooks, Lab Safety Training	
2		
3		
4		
5		
6	Background, Literature and Hypothesis Development	
7	Preparing Proteins	
8	Project Experiments & Data Analysis	
9	Project Experiments & Data Analysis	
10	Project Experiments & Data Analysis	
11	Conclusions and Presentation	
12		
13		
14		
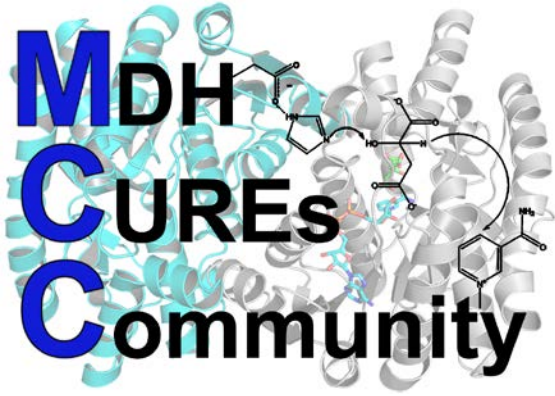
We Looked for Computational Approaches to provide data to address the questions we had identified

MDH CUREs Community

1 We needed **starting materials** to use

2 These were **prepared** and **characterized**

3 We asked questions about protein **conformation**, catalytic **activity** and ligand **binding**

4 We did some **preliminary experiments** with mutants to help refine our ideas of what we could/should do

5 We used web based **free** servers around the world to do experiments and **collect data**

6 We **compared wildtype** with **mutants**, and where possible with **literature data** for wildtype

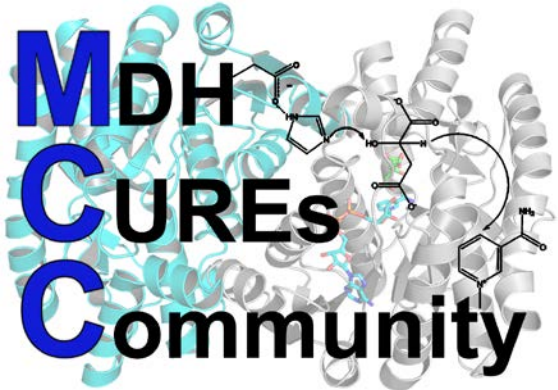**7** We collected **multiple sets** of all **data** for wildtype and mutants to allow **averaging** etc

8 We held mini-group meetings at the start and end of every lab period where **everyone had to talk**

For each computational Technique we used:

Journal Club to overview the computational approach and type of information it could give

We discussed the validity and limitations of the approach

We spent time at the end of each block of a particular computational approach discussing how to write the "method" and results etc

We usually set up computations at the end of a lab period to allow data to be ready to be analyzed at the start of the next lab

At the start of the next lab we walked through data analysis and then each student spent time analyzing their data

Each person, myself included, presented their data and we discussed analysis and conclusions

We walked through setting up the experiment on the server for one sample and then individually submitted the multiple jobs necessary to do the experiment in question

We talked about the wet lab experimental approaches that could be used to give complimentary data

Typical Flow of the Science: (and time required)
Real Time in Class, Prep for Next class
Protein Sequence:
     Bioinformatics, Clustal Omega etc
(https://weblogo.berkeley.edu/logo.cgi)
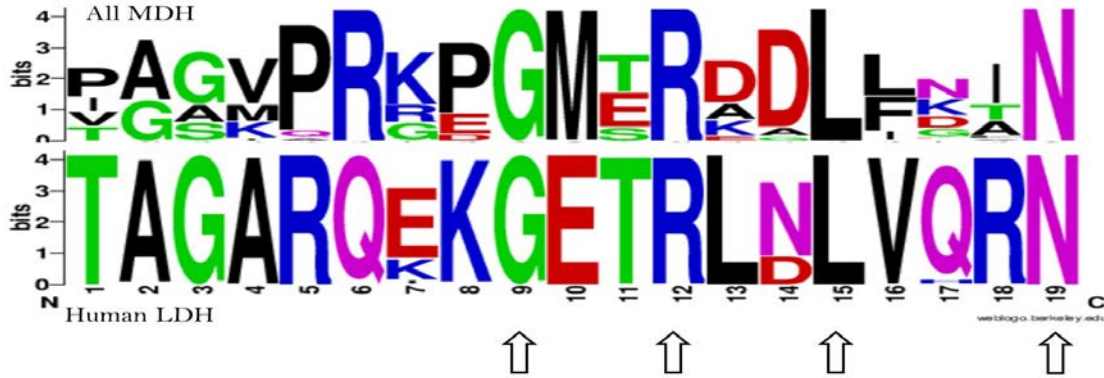
Make 3D Model
PyMol
Phyre2 Homology
Modeling
HawkDock

Refine Model
RefineD Energy
Minimization

⇧   ⇧   ⇧   ⇧

Ask Questions
Conformation/Local
Environment

Protein-Protein Interactions

Ligand Binding

Potential and Cryptic Sites

H++

HawkDock
MM/GBSA

SwissDock

POCASA
and
SwissDock

Lots of
Quantitative
Data!!!

Validate Model
MolProbity

Create Mutants as
appropriate to test
Hypothesis etc
**(No limit to how many
mutants you can make!!)**

Creating the Starting Material: Monomeric Proteins

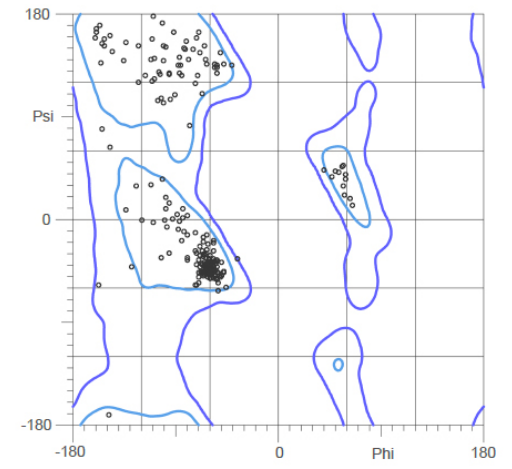**2 Sources**:

    1. Existing pdb file from the Protein DataBase (https://www.rcsb.org/)

    2. Model constructed from the amino acid sequence using Phyre2

(http://www.sbg.bio.ic.ac.uk/~phyre2/)

Making a Mutant:

2 Choices:

    1. Use PyMol to mutate the structure (https://pymol.org/2/)

    2. Use mutant Sequence in Phyre2

Refining the Resultant Structure

    RefineD (http://watson.cse.eng.auburn.edu/refineD/)
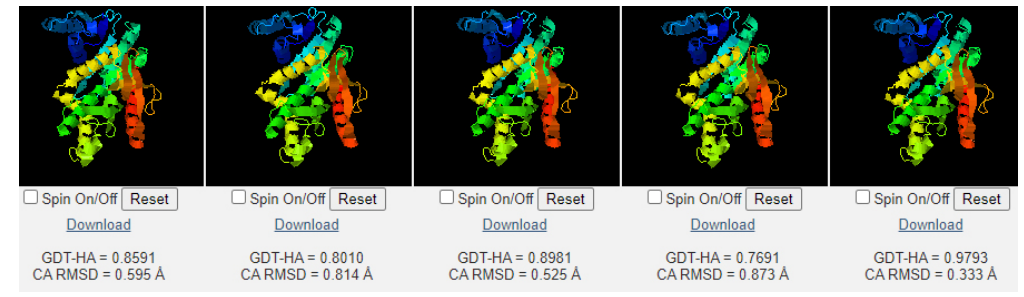
Characterizing the final Structures

    MolProbity (http://molprobity.biochem.duke.edu/)

Tutorials on setting up and submitting jobs to these servers can be found at:
https://mdh-cures-community.squarespace.com/virtual-cures-and-ures

Spin On/Off Reset	Spin On/Off Reset	Spin On/Off Reset	Spin On/Off Reset	Spin On/Off Reset
Download	Download	Download	Download	Download
GDT-HA = 0.8591 CA RMSD = 0.595 Å	GDT-HA = 0.8010 CA RMSD = 0.814 Å	GDT-HA = 0.8981 CA RMSD = 0.525 Å	GDT-HA = 0.7691 CA RMSD = 0.873 Å	GDT-HA = 0.9793 CA RMSD = 0.333 Å

MolProbity·Analysis·of·Subunits·of·Plasmodium·falciparum·Malate·Dehydrogenase:·
5nfr.pdb¶

Parameter¤	Average¤	Standard·Error¤	
¤	¤	¤	
Clash·Score¤	1.86¤	0.17¤	
MolProbity·Score¤	1.52¤	0.04¤	
¤	¤	¤	
Ramachandran·Residues·in·Favored·space¤	96.88¤	0.23¤	
Ramachandran·Residues·in·Allowed·space¤	99.85¤	0.09¤	

The·only·residue·found·as·an·"Outlier"·was·N276¶

Creating the Starting Material: Multimeric Proteins (homo-oligomeric Proteins, heteropolymers etc)

2 Approaches to obtain oligomers

    1. From Existing Crystal Structure using Pymol

    2. For proteins without a crystal structure or proteins with crystal structures but not of the complex using HawkDock (http://cadd.zju.edu.cn/hawkdock/)

First, make the monomers as previously, then assemble the oligomers
Can use constraints in HawkDock if you have them
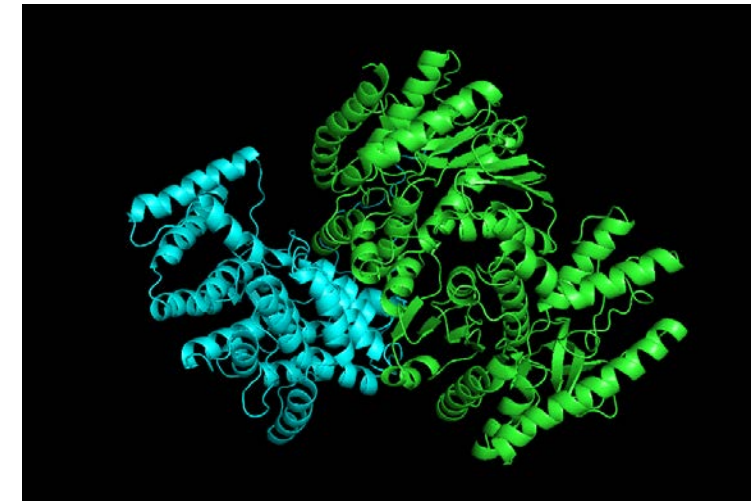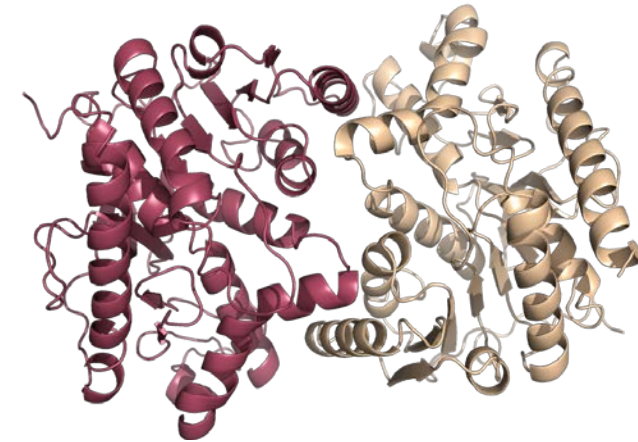
Tutorials on setting up and submitting jobs to create oligomeric proteins using these servers can be found at:
https://mdh-cures-community.squarespace.com/virtual-cures-and-ures

H++ (http://biophysics.cs.vt.edu/)

What Does it Do?

    Calculates pKa values from a structure

What You submit:

    pdb file (monomer, oligomer etc)

What You get back:

    i. Individual pKa values

    ii. What surrounding residues contribute to the altered pKa

What can you do with it?

    1. Explore impact of an interface in an oligomeric structure on pKa values(which pKa values are influenced by the interface etc)

    2. Relate changes in pKa values of mechanism related groups (eg catalytic base etc) to potential changes in activity (catalytic, binding etc)

    3. Compare Mutant and Wildtype protein with respect to local environment of titratable groups

Impact of Mutations at R130 on Loop and Catalytic Base pK values
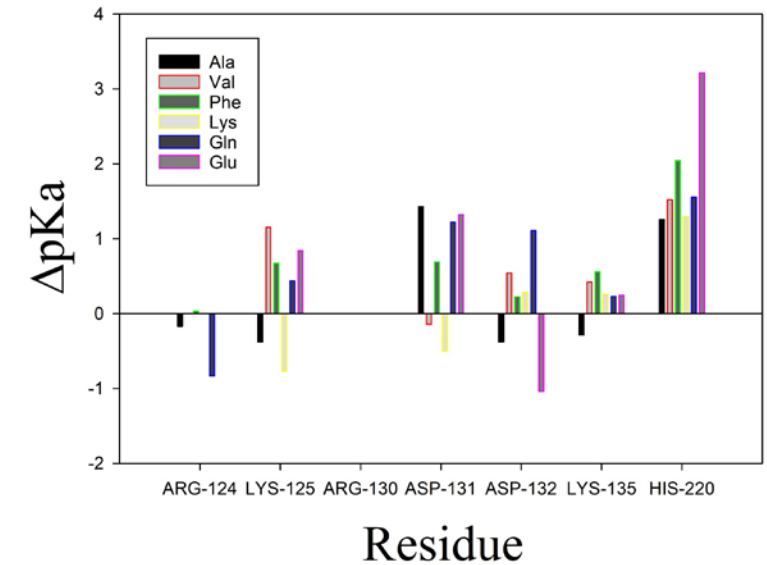
pKa values calculated from Phyre2 structures using H++
H++ Calculations: Salinity 0.15M, Local Dielectric 10
ΔpKa = Mutant-wt

Detailed Tutorials on setting up and using H++ can be found at:
https://mdh-cures-community.squarespace.com/virtual-cures-and-ures

Exploring Protein-Protein Interfaces with HawkDock MM/GBSA Module
(http://cadd.zju.edu.cn/hawkdock/)

What Does it Do?

Quantitates the contributions to Protein-Protein Interaction of every residue in each subunit in an Oligomer

What You submit:

The pdb file of the oligomer

What You get back:

Total energy of interaction between the designated subunits

Residue by residue contributions to the interaction energy- favorable and unfavorable

What You can Explore:

Nature of the interaction between two proteins

How mutations might change the interface

Compare Isoforms etc

Detailed Tutorials on setting up and using HawkDock can be found at:

https://mdh-cures-community.squarespace.com/virtual-cures-and-ures

Residue							
GLN-58	1.012	0.664		−0.07	−0.054	0.612	0.096
PRO-59	−1.968	−1.482		−0.07	0.01	−0.398	−0.198
MET-62	−2.762	−2.388		0.15	0.144	−0.25	0.674
LEU-63	−2.196	−1.27		−0.314	−0.06	−0.168	0.012
LYS-65	2.834	2.79		0.022	0.014	0.142	0.246
MET-66	−1.028	−1.306		0.348	0.098	0.404	0.086
PRO-82	−0.896	−0.862		−0.216	−0.154	−0.308	−0.08
GLY-83	−0.47	−0.628		−0.03	0.31	0.052	0.022
VAL-84	−0.948	−1.232		0.058	−0.172	−0.076	−0.286
ASP-87	−3.594	−2.57		−0.946	1.22	−1.656	−0.226
ILE-88	−1.196	−1.44		0.094	−0.098	−0.028	−0.346
HIS-90	−5.344	−3.568		−0.66	1.858	−0.394	0.98
MET-91	−3.286	−3.398		−0.252	0.088	−0.202	−0.252
ASP-92	1.798	1.906		0.314	0.142	0.558	−0.236
THR-93	−1.254	−0.97		−0.136	0.174	−0.062	−0.066
GLY-94	−1.832	−1.784		0.04	0.032	−0.082	−0.106
VAL-195	−1.63	−1.85		0.05	0.222	0.116	0.026
ARG-196	0.78	−0.042		0.928	0.706	1.582	−0.328
ASN-198	−4.884	−4.94		0.098	−0.86	0.208	−0.816
THR-199	−2.13	−2.668		0.484	−0.364	−0.132	−0.252
PHE-200	−0.594	−0.702		0.134	−0.128	−0.066	−0.036
PRO-209	−3.262	−3.542		0.008	−0.366	−0.088	−0.196
ARG-210	−0.868	−0.598		−0.102	0.134	−0.032	0.232
VAL-257	−3.596	−2.992		−0.16	−0.242	−0.756	0.03
LYS-261	−1.25	0.332		2.216	0.568	−1.09	1.612
ALA-267	−1.69	−2.322		0.85	−0.172	0.44	−0.306
THR-268	−2.164	−3.084		1.034	−0.618	0.564	−0.158
LEU-269	−5.024	−5.906		0.212	−0.056	−2.336	−3.556
SER-270	−2.714	−5.238		1.742	−1.468	1.88	−1.126
TYR-273	−3.156	−2.93		0.566	0.172	0.702	0.508
LYS-277	0	1.756		0.036	0.704	0.014	1.344

Detecting and Characterizing Pockets on a Protein using POCASA
(http://g6altair.sci.hokudai.ac.jp/g6/service/pocasa/)
Alone or in combination with SwissDock (http://www.swissdock.ch/)
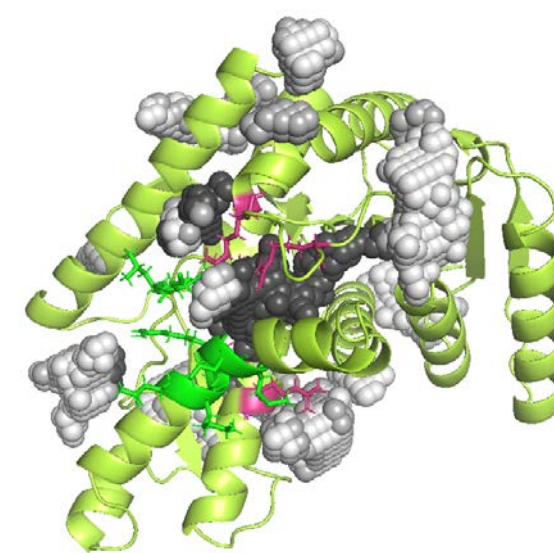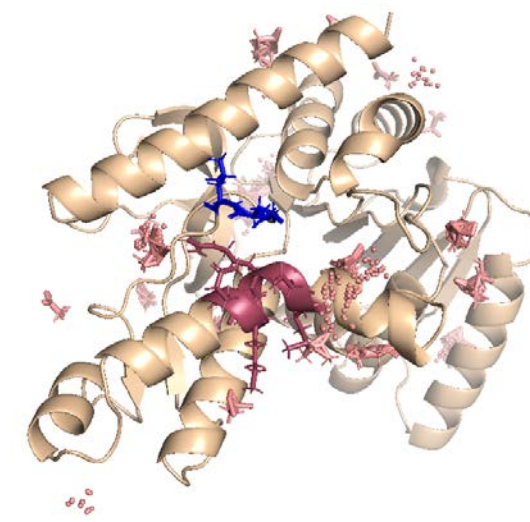
What They Do:

POCASA: Searches for cavities on the surface of the protein

SwissDock: Blind or Local docking of a designated ligand with a target protein

What you submit:

POCASA: pdb file of protein (monomer, oligomer etc)- You can use different size probes

SwissDock: pdb file of protein (monomer, oligomer etc) and Ligand structure (either supplied or selected from Zinc Database

What you get back:

POCASA: Location, area and depth information of cavities on the surface of the protein

SwissDock: Location and thermodynamic information on binding sites on the surface of the protein

What you can Explore:

Potential (Cryptic) Binding Sites on a protein including SAR relationships,

Physical Properties (in conjunction with PyMol electrostatics)

Impact of Mutations etc on sites on a protein

Isoform Differences etc

Tutorials on setting up and submitting jobs to these servers can be found at:

https://mdh-cures-community.squarespace.com/virtual-cures-and-ures

Exploring Specific Ligand Binding Sites with SwissDock

What it Does:

    SwissDock: Blind or Local docking of a designated ligand
with a target protein

What you submit:

    SwissDock: pdb file of protein (monomer, oligomer etc) and Ligand
structure (either supplied or selected from Zinc Database

What you get back:

    SwissDock: Location and thermodynamic information on binding
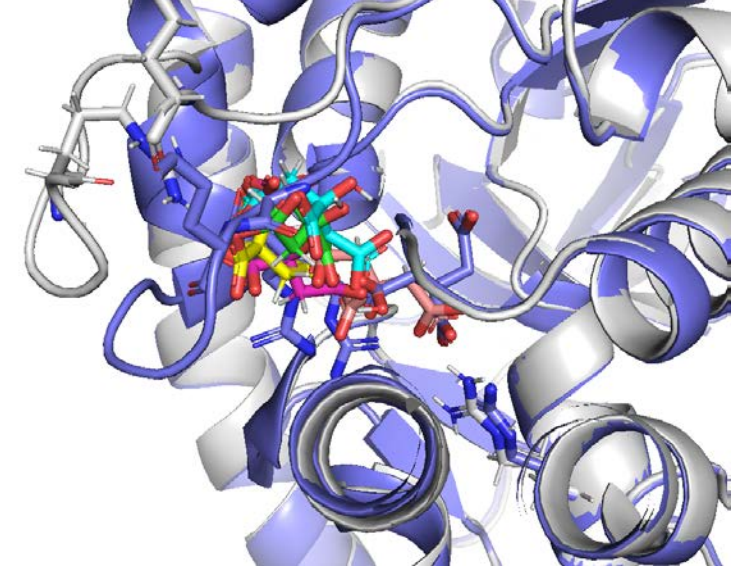sites on the surface of the protein

What you can Explore:

    Relative Affinity of Ligands for a Known Binding Site

    Thermodynamic Contributions to ΔG^o for a ligand

    SAR Relationships- Analogs, Inhibitors, Promiscuous Substrates etc

    Impact of Mutations etc on binding of a particular ligand to a
site on the protein. Isoform Differences etc

Detailed Tutorials on setting up and using SwissDock can be found at:
 https://mdh-cures-community.squarespace.com/virtual-cures-and-ures

pose		deltaG	deltaGelec	deltaGvdw	deltaGligsolvnon	deltaGligsolvpol	deltaGprotsolvnd	deltaGprotsolvpd	deltaGcompsolvr	deltaGcompsolvp	extraFull	surfFull	solvFull	IntraFull	InterFull	FullFitness	SimpleFitness	Energy
37.0000		-9.0635	0.0000	-122.9290	6.0283	-165.1190	215.3460	-1660.9500	214.5160	-1708.3500	0.0000	214.5160	-1708.3500	111.4400	-122.9290	-1505.3230	2.0128	2.0128
94.0000		-9.2652	0.0000	-136.7200	6.0996	-171.2790	215.3460	-1660.9500	213.9540	-1695.5100	0.0000	213.9540	-1695.5100	119.3280	-136.7200	-1498.9480	0.2701	0.2701
98.0000		-9.5004	0.0000	-138.9270	5.9588	-168.9470	215.3460	-1660.9500	213.7420	-1693.5000	0.0000	213.7420	-1693.5000	115.8240	-138.9270	-1502.8611	-0.8465	-0.8465
137.0000		-9.1179	0.0000	-129.9150	5.8989	-167.6970	215.3460	-1660.9500	213.9800	-1700.3900	0.0000	213.9800	-1700.3900	113.9970	-129.9150	-1502.3280	1.0820	1.0820
167.0000		-9.0020	0.0000	-135.4980	6.0661	-169.7000	215.3460	-1660.9500	213.8630	-1691.4300	0.0000	213.8630	-1691.4300	114.8580	-135.4980	-1498.2070	-1.5009	-1.5009
177.0000		-9.3386	0.0000	-138.9060	5.9324	-167.2520	215.3460	-1660.9500	214.0110	-1689.6500	0.0000	214.0110	-1689.6500	114.8300	-138.9060	-1499.7151	-2.2900	-2.2900
224.0000		-9.1427	0.0000	-139.2640	5.9736	-164.0040	215.3460	-1660.9500	213.4980	-1681.7900	0.0000	213.4980	-1681.7900	110.0170	-139.2640	-1497.5391	-3.3521	-3.3521

Papers for Journal Clubs on the Computational Techniques Presented today:

Kelley,L.A., Mezulis,S., Yates,, C.M., Wass M. & Sternberg, M.J.E. "The Phyre2 web portal for protein modeling, prediction and analysis" *Nature Protocols* volume 10, pages845–858(2015) https://pubmed.ncbi.nlm.nih.gov/25950237/

Bhattacharya D "refineD: improved protein structure refinement using machine learning based restrained relaxation".Bioinformatics. 2019 Sep 15;35(18):3320-3328. https://pubmed.ncbi.nlm.nih.gov/30759180/

Williams et al. (2018) MolProbity: More and better reference data for improved all-atom structure validation. Protein Science 27: 293-315. https://pubmed.ncbi.nlm.nih.gov/29067766/

Ramu Anandakrishnan, Boris Aguilar and Alexey V. Onufriev, "H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulation", Nucleic Acids Res., 40(W1):W537-541. (2012) https://pubmed.ncbi.nlm.nih.gov/22570416/

Aurélien Grosdidier,[1] Vincent Zoete,[1,*] and Olivier Michielin "SwissDock, a protein-small molecule docking web service based on EADock DSS" Nucleic Acids Res. 2011 Jul 1; 39(Web Server issue): W270–W277 https://pubmed.ncbi.nlm.nih.gov/21624888/

Gaoqi Weng, Ercheng Wang, Zhe Wang, Hui Liu, Feng Zhu, Dan Li, Tingjun Hou "HawkDock: a web server to predict and analyze the protein–protein complex based on computational docking and MM/GBSA" *Nucleic Acids Research*, Volume 47, Issue W1, (2019) https://pubmed.ncbi.nlm.nih.gov/31106357/

Yu J, Zhou Y, Tanaka I, Yao M "Roll: a new algorithm for the detection of protein pockets and cavities with a rolling probe sphere". Bioinformatics. 2010 Jan 1;26(1):46-52. https://pubmed.ncbi.nlm.nih.gov/19846440/

Working as Part of a bigger project: eg MCC Malate Dehydrogenase projects

Join an Ongoing Multi-Institution CURE Project
(if you are interested, email me: **jbell@sandiego.edu**)

1. The Active Site Loop Project
2. The Subunit Interface Project

On going Goals:
Computational work to support existing and future wet lab work

Aims:

1. Answer important Scientific Questions about Malate Dehydrogenase
2. Write and Publish papers involving many student and faculty coauthors
3. Provide preliminary results for potential future grant applications

Thankyou: 1. ASBMB and the Student Chapters for putting this together
2. All of You
3. NSF-1726932 EHR-IUSE

Join us for Hands on Training on all the computational Techniques presented here:

Thursday August 6th
Friday August 7th.

Email:
molecularlifescience@gmail.com
& we'll send you the zoom links

Dr Tamara Mans

Building in Collaboration

Dr Betsy Martinez-Vaz

Adapting to any protein you want to use:

Dr Anthony Bell