

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase

## Construction & Expression Module

Malate Dehydrogenase Laboratories	Page
Overview	2
Learning Goal: <i>Work Safely &amp; Keep an Accurate Lab NoteBook</i>	2
Maintaining an Accurate Lab NoteBook	2
<i>Rubric for Learning Goal</i>	5
Learning Goal: <i>Successfully follow &amp; trouble-shoot a protocol, explain the basis for the steps involved and understand the variables that must be controlled for</i>	6
Using Protocols	8
Isolating Plasmid DNA	8
Primer Design	8
Site Directed Mutagenesis	11
DNA Sequencing	11
Archiving the Cells/DNA	11
Expression & Purification of a His Tagged Protein	12
<i>Rubric for Learning Goal</i>	13
Learning Goal: <i>Be able to obtain and use quantitative data and appreciate the need for reproducibility and appropriate statistical analysis</i>	14
Purification of a His Tagged Protein	14
Specific Activity Determination	14
Molecular Weights & Purity	14
SDS PAGE	17
Native PAGE	17
Size Exclusion Chromatography	17
Western Blotting	17
Averages, Errors, Confidence Limits and Quantitative Calculations	17
<i>Rubric for Learning Goal</i>	30

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase

## Construction & Expression Module

### Overview

*The construction and expression module, in addition to constructing the materials that are an integral part of your project, has three specific learning goals:*

1. [Work Safely & Keep an accurate lab notebook.](#)
2. Successfully follow a protocol, explain the basis for the steps involved and understand the variables that must be controlled for.
3. Be able to obtain and use quantitative data and appreciate the need for reproducibility and appropriate statistical analysis.

**Learning Goal:** [Work Safely & Keep an accurate lab notebook.](#)

### Maintaining an Accurate Lab Notebook

*Why have a laboratory notebook? Isn't this just busy work? Don't worry; I'll remember what I did in lab. I'll just write down my results on scratch paper and fill in the details later.* We have all made these comments about laboratory notebooks. Lab notebooks are not the students (nor the professors) favorite part of laboratory or research classes. However, communication is an essential part of conducting laboratory experiments. We gain practice in keeping notebooks in classes so we know how to do it when we work in research laboratories. Much of the value of collecting experimental data is lost if proper records are not kept. If methods and results are not clearly written and recorded, this information cannot be effectively transmitted to other scientists (including your professor). Notebooks contain your data whether good or bad, observations from your experiments, and form the basis for every scientific paper you write. Notebooks are legal records for documenting drugs, biologics and medical device research under FDA guidelines. **The following guidelines should be strictly adhered to for the recording of laboratory exercises.**

Your laboratory notebook should be an accurate record of what you do in the lab, and should contain notes and calculations as well as appropriate comments regarding the experiments.

***A major function of a lab notebook is to allow another competent scientist to reproduce EXACTLY your experiment.***

The course uses an electronic lab notebook (ELN). You have received an email with a link to access the notebook and register for the LabArchives (host of our ELN). Your notebook will be graded based on the formatting (shown below), completeness and the statement in bold above. No grade will be assigned for neatness, but a disorganized lab notebook will be penalized. You will have 3 notebook checks throughout the semester. Your notebook will also be checked prior to lab (14x) to ensure that you have prepared for that day's experiments.

Finally, when taking notes for the lab (lecture or self made notes), these should each have their own entry in the lab book.

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

## A. Lab Blocks

- The ELN template is organized into the 6 lab blocks. Within each block, a lab page is assigned to each day's experiments. Headings and a rich text box have already been added to each page. You will edit the rich text box and add additional elements to document your experiments, results and conclusions.

## B. Date

- Done automatically! The ELN time stamps all entries using the National Institute of Standards and Technology standard time.

## C. Title

- Headings have been added to each page. You may edit these or add additional title bars, but use an intelligent title, not just the name of the experiment copied from the laboratory manual. Pick a title that describes the content of your work. Particularly important for multi-day experiments.

## D. Introduction/Purpose

- This section should include a two or three sentence statement of the purpose(s) and objective(s) of the exercises being performed.

## E. Materials / Procedure / Protocol or Methods

- Write a description of procedures used including any deviations from the information presented in the laboratory handout.
- **If there is a published exact protocol, then reference the protocol or procedure AND include a simple outline, flow chart or description of the referenced protocol. DO NOT copy & paste the protocol into the notebook.**
- When you have to repeat a routine operation several times, such as a protein assay, you can skip the details and link to the specific page in the ELN where the complete procedure was originally described.

**The title, introduction/purpose and materials and methods sections should be completed before coming to lab.** (Deviations from the information presented in the laboratory handout will be recorded as you perform the laboratory experiments.)

## F. Data and Results – Including Analysis

- All data and observations that are generated should be recorded in your laboratory notebook at the time of the exercise. If you upload a photo, you **MUST** include a figure title and legend to explain the significance of this data. This should include any tables, graphs, formulas or other information from the laboratory manual. Your data analysis should be performed as soon after the conclusion of an experimental procedure as possible and before the next lab period. Any tables of the data collected that day and graphs of that data should be included if appropriate. This section should include all calculations, averages and corrections to the recorded data.
- All information should be neatly presented with graphs and tables labeled appropriately. Graphs can be prepared using iPad apps or your own computer and uploaded to the appropriate ELN page.

## G. Discussion and Conclusions

- This section should include any interpretations, conclusions, or suggestions regarding the results of that day's exercise. A discussion of the expected results and why they were or were not obtained

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

should be included. THIS IS NOT A SUMMARY OF EVERYTHING THAT HAPPENED DURING THE EXPERIMENT THAT DAY. THIS IS A DISCUSSION OF THE DATA AND FINAL RESULTS.

- A good discussion might include:
  1. What were the major points illustrated by the data?
  2. Do the results agree with previously published works?
  3. Is the data contradictory in itself?
  4. Does your research have potential for follow-up experiments?
  5. Do your results support or disprove your hypothesis?
  6. Are your results dramatically different than what was anticipated and if so why?

## H. References

- Include any references that were consulted for the experiment or cited in the report. Minimally, this should include your laboratory manual.
- References should be presented in alphabetical order by the last name of the first author.

## Other Notes:

1. For this lab, you will use the electronic notebook from LabArchives. If you have any questions regarding access to this ELN ask your instructor.
2. Since this is a course, taking notes in the laboratory notebook is allowed. Simply place these in a rich text box on the relevant ELN page, including a header above the notes with a descriptive title.
3. Number and label all tables and graphs with a title indicating what they are intended to represent.
  - A. Tables are numbered and titled above the table.
  - B. Figures are numbered and titled below the figure.
4. This is not a personal diary and references to that effect should not be included.
5. Show your calculations in the data/results section. Often if an experiment does not work it can be tracked down to a miscalculation.
6. Remember your results from the exercise are important. You are not being graded on the outcome of experiments therefore you should strive for the most accurate and intelligent representation of your data possible.
7. **Copy and pasting printed methods from the web does not count as a written protocol.** You can certainly refer to it, but do NOT just copy and paste.

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase

## Construction & Expression Module

### Rubric for Learning Goal 1:

<i>Criteria</i>	<i>Excellent</i>	<i>Acceptable</i>	<i>Fair</i>	<i>Unacceptable</i>
General Clarity Required elements present and notebook current	Writing and organization allows reader to easily follow flow, understand and find information. All required elements present and up-to-date at submission.	Writing and organization allows reader to follow flow, understand and find information with a small amount of effort. All items present but some elements (e.g. data analysis/interpretation) incomplete for most recent expt (within 1-2 lab periods of submission* <sup>2</sup> ).	Writing not always clear, or poorer organization requires more effort from reader to follow flow, understand and find information. Significant items absent from experiments performed within 2 – 3 lab periods of submission* <sup>2</sup> .	Writing not clear or organized. Substantial required elements absent/incomplete for expts completed more than a week ago.
Safety	Appropriate PPE noted and safe handling and disposal noted as appropriate, refers to MSDS	Appropriate PPE noted and safe handling and disposal noted as appropriate	Fails to mention one of appropriate PPE or Safe handling or disposal as appropriate	Fails to mention two of appropriate PPE or Safe handling or disposal as appropriate
Procedures	Clearly and concisely describe steps of experiment, showing set up calculations and preps (e.g. to make solutions) and all necessary information to repeat.	Some information absent or incomplete but generally possible for a trained biochemist to attempt to repeat expt.	Significant omission of information making it difficult for a trained biochemist to repeat expt with certainty.	A trained biochemist could not repeat the expt.
Data/Observations	Complete, related to procedures and necessary information given (e.g. units). Good organization of data in tables or other ways to present clearly. Tables/Graph embedded or included as attachments.	Complete, but not always related to procedures or some necessary information absent (e.g. units). Tables/Graph embedded or included as attachments.	Not complete (e.g. Figures not embedded or not included as attachments.) or not always related to procedures or significant necessary information absent (e.g. units).	Significant data and observations or necessary information absent.
Results (Data Analysis)	Calculations and analysis using data are complete and logical. Figures, tables, or other graphics <i>clearly</i> present results or data with titles and figure legends that add appropriate context and necessary expt parameters.	Calculations and analysis complete and logical. Figures, tables, or other graphics <i>clearly</i> present results or data with titles and figure legends that add appropriate context but lack some minor information necessary to interpret them.	Calculations or analysis not complete. Figures, tables or other graphics present results or data but contain errors or lack significant information necessary to interpret them.	Necessary calculation or analysis absent. Data not presented as figures/tables or other graphics where appropriate.
Discussion/ Conclusion (Data Interpretation)	Logical conclusion(s) based on goals; Methods, observations and interpretations are evaluated and placed into context; thoughts and consideration of broader consequences of result(s).	Logical conclusion(s) based on goals; Methods, observations and interpretations are evaluated but not placed in context or less consideration of broader consequences of result(s).	Only portions of the results are evaluated, or contain unsupported assertions or less attempt to consider broader consequences.	No evaluation or interpretations of results are presented.

# **MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase**

## Construction & Expression Module

*Learning Goal: Successfully follow & trouble-shoot a protocol, explain the basis for the steps involved and understand the variables that must be controlled for.*

### **Overall Rationale to a Site Directed Mutagenesis Exploration of Structure-Function Relationships**

As you will see, altering specific amino acid residues in the primary sequence of a protein is quite a simple process. However, to effectively use the approach of site directed mutagenesis you need three dimensional structural information. The availability of sequence homology information can also be of significant help in designing appropriate mutants. There are two appendices to this laboratory which detail how you can obtain and analyze both sequence information and structural information. If you are not familiar with these protocols you should read these appendices.

Once you have made the mutants there are still a significant number of problems that you might encounter and while these will be discussed in more detail in more advanced laboratories it is useful to be aware of them at this point.

First, just because you design and make a “mutant” there is no reason to know that the mutant will be expressed and fold correctly- the residue mutated may have played some role in the folding process that you had no reason from the final three dimensional structure to be aware of. In such a case the protein will be expressed, will not fold correctly and is likely to be rapidly degraded.

Second, to probe the effects of the mutation you must first purify the protein and be able to characterize properties that will likely have been affected by the mutation. Enzyme kinetics studies can give a great deal of information about the functional properties of a protein, and as you design a mutant you should be thinking about what kinetic parameters you might measure to follow the effects of a mutant. For example if you choose to alter a residue that you speculate plays a role in catalysis you would want to measure the maximum rate of the enzyme and know that it is due to catalysis and not due to a product release step. If you change a residue that may play a role in substrate binding you must be able to determine through either initial rate kinetics or direct substrate binding studies that you have an alteration in affinity for the substrate. If you change a residue that may affect the quaternary structure of the protein and affect perhaps subunit cooperativity in an allosteric enzyme you need some parameter that tells you about the cooperativity and the strength of the subunit interactions.

Finally, you should be aware that even if a protein folds and is functional at some level, alterations in the function could have resulted in some overall change in shape triggered by the site directed mutant and not reflect a direct role of the particular side chain in the reaction or substrate binding etc. To guard against such a possibility you need to be able to measure some parameters of the overall shape and conformation of the protein and show that there have been no gross changes in structure. The ultimate proof that the conformation of the protein has not changed is of course to crystallize the protein and determine the three dimensional structure of the mutant.

To explore structure-function relationships in the enzyme you can use a variety of bioinformatic techniques and site directed mutagenesis. To do this you need to purify the expression plasmid that was used to express the

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

native protein you purified earlier and you will start the process of site directed mutagenesis by designing the mutant that you wish to make. Subsequently you will isolate the mutant DNA and confirm the sequence (and mutation) by DNA sequencing. Since the mutants that you will make have never been made before it is essential that you design the primers and let your lab instructor know what they are this week so that they can be ordered. The design of the primers is described in complete detail in the “QuikChange Site Directed Mutagenesis Manual” that is part of this laboratory module. You will make the mutants using a thermocycling approach and place the DNA that you create into an expression cell.

Shown below is the nucleotide sequence of the cDNA clone of the precursor of watermelon glyoxasomal malate dehydrogenase:

```
1      caacgctaag ttcccaaagg tttctgatct tgaagcgggtt ggtttgtttt tctgtttgct
61     aaactaatta tgcagccgat tccggatggt aaccagcgcgca ttgctcgaat ctctgcgcat
121    cttcatcctc ccaagtctca gatggaggag agttcagctt tgaggagggc gaattgccgg
181    gctaaaggcg gagctcccgg gttcaaagtc gcaataacttg gcgctgccgg tggcattggc
241    cagccccctt cgatggtgat gaagatgaat cctctggttt ctgttctaca tctatatgat
301    gtagtcaatg cccctgggtg caccgctgat attagccaca tggacacggg tgcctggtg
361    cgtggattct tggggcagca gcagctggag gctgcgctta ctggcatgga tcttattata
421    gtcctctcag gtgttctctg aaaaccagga atgacgaggg atgatctggt caaaataaac
481    gcaggaattg tcaagactct gtgtgaaggg attgcaaagt gttgtccaag agccattgct
541    aacctgatca gtaatcctgt gaactccacc gtgcccatcg cagctgaagt tttcaagaag
601    gctggaactt atgatccaaa gcgacttctg ggagttacaa tgctcgacgt agtcagagcc
661    aatacctttg tggcagaagt attgggtctt gatcctcggg atggtgatgt tccagttggt
721    ggcggctcatg ctggtgtaac cattttgccc cttctatctc aggtgaagcc tccaagttct
781    ttcacacaag aagagattag ttacctgact gataggattc aaaatggtgg aacagaagtt
841    gtcgaggcca aagcaggagc tggtcagca actctctcaa tggttatgc tgccgttaag
901    tttgcagatg catgcctcag gggcttaaga ggagatgctg gtgtcattga atgcgcggtt
961    gtgtcttctc aggtgactga acttccattc tttgcatcaa aagtacgact tggctgcaat
1021   ggtatcgaag aagtatactc ccttggcccg ctaaagatg atgagaggat tggattggag
1081   aaagcgaaga aagagttggc aggaagcatt gagaagggag tttccttcat cagaagctga
1141   agagatgcca attaccatta gttttaatag aacattcca tctcttatag attacttgtg
1201   ctcaatgttt tcttgagat tgaagttgat tgaatgata ccacaccacg tattttata
1261   ctaataaaac tatatcgcca tcatgtcgat atttaatgca caaccaaag ggttggatta
1321   gagtaccttt tatg
```

Together with the protein sequence: (the precursor sequence which is removed upon organelle import is shown in green)

```
1      mqpipdvnqr iarisahllp pksqmeessa lrrancrakg gapgfkvail gaaggigqpl
61     amlmkmnplv svlhlydvvn apgvtadish mdtgavvrgf lgqqqleaal tgmqliivpa
121    gvprkpgmtr ddlfkinagi vktlcegiak ccprainvli snpvnstvpi aaevfkkagt
181    ydpkrllgvt mldvvrantf vaevlgl DPR dvdvpvvggh agvtilppls qvkppssftq
241    eeisyltdri qnggtevvea kagagsatls mayaavkfad aclrglrgda gviecafvs
301    qvtelpffas kvrlgrngie evyslglplne yeriglekak kelagsiekv vsfirs
```

How does one go about finding the protein sequence from the nucleotide sequence?

In this case life is simple: the first atg codon (**shown in bold**) is the initial methionine in the precursor sequence. With many cDNA sequences however it is not that simple and to translate the nucleotide sequence often means trying a number of atg codons until the correct reading frame is obtained: often the reading frame set by the first atg codon is not correct and the sequence will not translate completely: you run into stop codons well before the

# **MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase**

## **Construction & Expression Module**

correct stop codon and get a prematurely terminated sequence. Knowing where the sequence stops is another potential problem: when? There can be three stop codons, uaa, uag and uga, and one of them, UGA, sometimes codes for selenocysteine (requires the presence of a cis-acting selenocysteine insertion sequence (SECIS) element. Similarly, it is thought that Pyl (Pyrrolysine) is inserted by UAG codons with the help of a putative pyrrolysine insertion sequence (PYLIS) element) and hence there could be a problem of knowing where the reading frame actually stops. In the case of the watermelon glyoxasomal malate dehydrogenase you will find that the actual stop codon is a uga codon and that the next stop codon is some six residues later (uag). If the potential confusion involves a relatively long stretch of amino acids the molecular weight of the protein can be useful in deciding the actual stop codon.

### **Using Protocols**

Many procedures used for purifying DNA, Site Directed Mutagenesis, Protein Expression and Purification use commercially available kits and protocols. The Handbooks for these kits usually contain a variety of protocols that can be adapted to a particular situation. The ability to read and quickly discern what are the relevant sections is a valuable skill.

### **Purification of the Watermelon Glyoxasomal MDH Expression Plasmid**

You will use the Qiagen Plasmid Miniprep spin kits to isolate the plasmid DNA- read and follow the provided protocol from the Company

### **Quantitation of the DNA**

**DNA in solution at 260nm has an absorbance of 1.0 for a 50microgram/mL solution. This allows you to quantitate the DNA that you have purified by measuring its absorbance at 260nm.**

**You should record a uv spectrum of your isolated DNA, using 240-300nm: this can be done using the same solvent that you used to elute the DNA in the final stage of the miniprep. Blank your spectrophotometer and record a baseline spectrum, then add approximately half of your isolated DNA and record the spectrum. Use 0.75mL of solvent to blank your spectrophotometer.**

### **Designing Primers**

**The principles and protocol for designing primers for use with QuikChange mutagenesis are described in the Qiagen Handbook**

### **Designing the Primers using BioEdit**

During this lab period you will design the DNA primers to alter the watermelon glyoxasomal malate dehydrogenase gene.

To begin, what mutation(s) have you selected to create? Provide the mutant in the following format: single letter amino acid code of wild type sequence residue # single letter amino acid code of mutated sequence. For example if mutating the 51st residue of the MDH sequence from a glycine to a proline: G51P

Describe the basis for this mutation. For example, how does the mutated residue alter the functional group available to the enzyme? Have you altered the surface area occupied by the amino acid side chain?

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

Copy the DNA sequence of wgMDH provided in your ELN for Block 4 into the ClipBoard. Open BioEdit. Under File>New from Clipboard.

Rename the DNA sequence from ~out to WT wgMDH by double clicking on the current name. Translate the DNA sequence into the protein sequence: Select the sequence by clicking on “WT wgMDH” >Sequence>Nucleic Acid>Translate>Frame 1.

You now have numbered nucleotide and protein sequences for your gene for the top strand of DNA. Go to the protein residue(s) that you will mutate. What is the number of the nucleotide encoding the first base of the codon for this amino acid?

Create the bottom strand or complement of your WT wgMDH sequence: Select your WT wgMDH sequence: >Edit>Copy Sequence. Unselect your WT wgMDH sequence. >Edit>Paste Sequence. Rename your top strand: WT wgMDH 5. Select the copy of the WT wgMDH sequence: >Sequence>Nucleic Acid>Complement. Rename the complement sequence: WT wgMDH 3’

Copy the 45 nucleotides (ntes) centered around the site of your mutation: Hold left mouse button & drag over 5’ and 3’ sequence to select, Ctl-C to copy. Paste into NEW Rich Text box in your ELN. Label the ends of your sequence indicated 5’ or 3’ and nte number. Change the font to Courier New. Making life easier: Place a space between every 3rd base to delineate codons. Above the 5’ strand enter the 1 letter AA code (Remember, you have a nte/protein sequence in BioEdit). Highlight the codon that you will alter by changing the font color. If you are inserting or deleting a sequence, highlight the AA codons bounding the insertion/deletion.

Example: Mutation at Residue 8: N8L

M Q P I P D V N Q R I A R I S

5’ 1-ATG CAG CCG ATT CCG GAT GTT AAC CAG CGC ATT GCT CGA ATC TCT-45 3’  
3’-TAA GGC CTA CAA GAC GTC GCG TAA CG- 5’

5’ ATT CGG GAT GTT CTG CAG CGC ATT GC- 3’  
3’1-TAC GTC GGC TAA GGC CTA CAA TTG GTC GCG TAA CGA GCT TAG AGA-45 5’

Identify codon that you will use to replace wild type codon.

Go to [www.kazusa.or.jp/codon/](http://www.kazusa.or.jp/codon/)

Query box: Escherichia coli (Case: insensitive) >Submit. This will bring up a list of E. coli genomes. Choose whichever you like. You will be given a codon usage table but it does not include the amino acids for which the codon codes.... Bummer!

Under Format: Select a Code = Eubacterial; Codon Usage Table with Amino Acids; >Submit. AAhhh, much better.

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

Now you have a codon usage table with the 1 letter amino acid codes as well as the frequency of codon usage in *E. coli*.

Why select *E. coli* and not another classification?

Identify the codon that corresponds to your mutation and has the highest frequency in *E. coli*. For example, leucine has 6 codons with CUG having the highest frequency (0.49).

Designing a primer.

Keep in mind the rules outlined in the QuickChange Lightning Site-Directed Mutagenesis Kit manual (pgs. 9-10).

In your ELN entry, add a couple of carriage returns between your 5' and 3' DNA sequences.

For your primer annealing to the 5' strand, type it below this strand. For your primer annealing to the 3' strand, type it above this strand.

Highlight your mutated codon by color. Using the formula in the QLSDM manual, calculate your  $T_m$ . What is your percentage of GC content?

Use the program OligoCalc to predict  $T_m$ , hairpin and loop formation of your primers.

<http://biotools.nubic.northwestern.edu/OligoCalc.html>

Copy your primer sequence into the Nucleotide base codes box: >Calculate.

What information are you given? Does this agree with your calculations? Note that you can alter the primer and salt concentration. What effect does changing these parameters have on your results?

Go to [www.agilent.com/genomics/qcpd](http://www.agilent.com/genomics/qcpd)

Follow their instructions for primer design. You will paste the DNA sequence found in your ELN.

Compare the primers that you designed to those that were predicted by the program.

***Provide to your instructor before the end of lab:***

Mutant: \_\_\_\_\_

5' Primer sequence written in 5' to 3' direction (This is extremely important!!!)

$T_m$  of 5' primer \_\_\_\_\_

GC content of 5' primer \_\_\_\_\_

3' Primer sequence written in 5' to 3' direction:

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase

## Construction & Expression Module

Tm of 3' primer \_\_\_\_\_

GC content of 3' primer \_\_\_\_\_

### Site Directed Mutagenesis

The protocol for creating a site directed mutant by QuikChange mutagenesis are described in the Agilent Handbook

### DNA Sequencing:

Prepare a sample of the mutant DNA for sequencing by Eton Biosciences

### Analyzing sequencing results:

Download the .abi and .gb files from Bb>Lab blocks>Experiment 4

Open both files in BioEdit.

Copy wgMDH DNA sequence: Ctl-F8 or >Edit>Copy Sequence.

Paste wgMDH DNA sequence into HA5IKEA.CMV... window: Ctl-F9 or >Edit>Paste Sequence.

Highlight both sequence names: Hold left mouse button drag over both names until black highlighted.

Align two sequences: >Sequence>Pairwise alignment>Align two sequences (allow two ends to slide). This should open two new windows: one will show sequence alignment, the other the two aligned sequences.

Scroll through the aligned sequences. Note sequence identity and where sequences differ.

Identifying areas of differences: Select wgMDH DNA sequence. Drag this sequence so that it is listed first. Now, select the two sequences by selecting both names. >Alignment>Plot identities to first sequence with dot. A new window will appear with this alignment. Scroll through the sequences. What do you see and what does it tell you? Note the first place where you observe a mismatch. What nucleotide is this in the HA5IKEA.CMV sequence? Hint – click on the nucleotide, the nte # then appears in the window information. What mutation was incorporated into the DNA? Identify all mutations inserted into the wgMDH DNA.

Go to the ABI Chromatogram. Scroll to the nte that you identified as the first mutation. From the chromatogram, is this data consistent with the reported sequence for the mutant? Scroll through the chromatogram. At what point can you no longer discern the sequence? Given this information are you confident in the sequence that pertains to the MDH gene in your alignment.

### Archiving the Mutant Cells

Once you have plated the Quikchange products they should sit in the 37°C incubator no longer than 16-24 hours depending upon the size of the colonies. At that point you should start about 4-5 “starter culture” tubes from a colonies from the plate and make 2 tubes of “Glycerol Stocks” from each starter colony. These glycerol stocks should be clearly labeled and frozen in the –80m freezer for subsequent use.

# **MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase**

## **Construction & Expression Module**

### ***PREPARATION OF GLYCEROL STOCKS:***

#### ***Prelabel autoclaved storage tubes***

***To autoclaved [ie sterile] tubes add 1mL of the cell suspension from the starter culture –make sure that the cells have not settled before pipeting the 1mL.***

***Using specially prepared “Blue” tips [these tips have had a little of the end removed to make it easier for them to be used with the viscous glycerol solution prior to autoclaving] add 0.25mL of autoclaved, sterile, glycerol.***

***Vortex briefly***

***Place in storage rack: label storage rack and place in –80 Freeze***

#### ***Expression of your Mutant Protein***

# **MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase**

## Construction & Expression Module

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase

## Construction & Expression Module

<i>Learning Objective</i>	<i>Excellent</i>	<i>Acce</i>	<i>Poor</i>	<i>Unacceptable</i>
Is aware of the nature of an expression plasmid and its use in site directed mutagenesis	Can use “snappene” viewer and describe key features of the appropriate expression plasmid, for example the pQE60 expression plasmid used with wt-WgMDH			
Can find and follow protocols to isolate, quantitate and mutate plasmid DNA. Can obtain DNA sequencing data and analyze using SnapGene or Bioedit	Given a task and standard commercial protocols (for example Qiagen plasmid miniprep and quickchange) can identify and use the appropriate protocol, and can explain the basis for the steps involved in the protocol. Can prepare suitable samples for DNA sequencing.			
Can find, describe and use protocols for an affinity resin to purify an appropriately tagged protein	Uses the Qiagen “Expressionist” handbook to select the appropriate protocol for purification of His Tagged wt-WgMDH or mutant forms. Understands how expression can be monitored and optimized			
Can describe and use both uv spectroscopy & the Bradford Dye binding assay protocols to determine the amount of protein in a solution and is aware of the limitations of both.	Describes the physical basis of both uv and Bradford approaches to protein quantitation. Understands the role of the “extinction coefficient” in uv approaches and the necessity of a standard curve with Bradford approach. Describes the assumption about the suitability of the standard protein			
Can describe and use both continuous and stop-time assays to determine the initial rate of an enzyme catalyzed reaction and is aware of the limitations of both approaches	Describes the basis of continuous and stop time rate assays and the importance of the rate being linear and extrapolating back to the starting Absorbance. Describes suitable controls and “blanks”. Understands the importance of controlling a variety of variables.			
Can determine the specific activity of an enzyme and use this information to quantitate the amount of enzyme in a solution	From appropriate data can accurately calculate the specific activity of a sample, and given the specific activity of the pure protein can calculate the amount of enzyme in a given sample. Can relate specific activity and turnover number for a pure protein.			
Understands that in protocols such as described here there are variables that must be controlled for and can trouble shoot the protocols	Can identify variables that must be controlled for a protocol to be successful. Can qualitatively predict the impact of uncontrolled variables on the success of a protocol			

# **MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase**

## Construction & Expression Module

### **Expression and Purification of a His Tagged Protein II**

***Learning Goal:** Be able to obtain and use quantitative data and appreciate the need for reproducibility and appropriate statistical analysis*

**You will use the protocols described in the Expressionist Handbook to purify the mutant protein**

#### **Initial Characterization of the Activity**

**Goal:** You have performed an expression and purification of wt-wgMDH. Now you need to characterize the enzyme in each fraction and the overall impact of each expression condition. To do this, you will assay the protein concentration for each fraction, observe the quality of the purification procedure, and measure the enzyme activity in your fractions.

#### **Specific Activity Protocol**

#### **Molecular Weight Determination**

##### **Introduction**

In the previous Sections we discussed the purification of proteins using their size/shape parameters and techniques using adsorption chromatography based on either their native charge or hydrophobicity or biospecific affinity. By using a variety of combinations of these techniques it is usually possible to purify a protein sufficiently for further study. Often then it is necessary to a) establish the purity of the “purified” protein and to characterize its molecular weight. In this laboratory we will explore the various techniques that will tell us directly about the charge, size and shape of proteins. Since proteins tend to have unique charge and size properties this discussion will also lead us into a consideration of the criteria of purity of a protein, and thus is a logical conclusion to our discussion of the purification of proteins. Since the charge and size of a protein is directly related to its amino acid sequence we will also consider various aspects of "proteomics", the analysis of protein structure and function in terms of the expression of proteins in a particular cell type: how can you experimentally determine what proteins are expressed at a given point in time in a given cell line? As you will see the techniques that give information about molecular weight, shape, charge and purity of proteins can give a wealth of information in the arena of proteomics.

While you may not have time to sample each of the techniques described here it is important to approach these experiments with an understanding of several basic principles concerning molecular weights and basic protein structure.

##### **Overview of Molecular Weights and Molecular Weight Determination Methods**

The molecular weight of a protein polypeptide chain [the smallest covalently bonded unit that a protein can be broken down into] can give you information about a protein that cannot be obtained from the gene sequence. While the gene sequence gives you a region of contiguous peptide bonds it does not tell the complete story

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase

## Construction & Expression Module

about a protein structure: it does not tell you about any post translational modification events that may have taken place [proteolytic processing, phosphorylation, glycosylation, lipidation etc], nor does it tell you about the native structure of the protein: does it contain subunits, either of identical sequence [homopolymer] or of different sequence [heteropolymer], nor does it tell you about interchain disulfide bonds. Much of this information [or in the case of post translational modification- at least a clue to its existence] can be obtained by careful application of various molecular weight determination techniques.

Furthermore, a consideration of the “type” of molecular weight that a particular method gives can lead to clues about the overall purity of a sample. What types of molecular weights are there?

### Number Average Molecular Weights

A “Number” average molecular weight is obtained when the experimental technique used depends upon the number of particles present, ie a technique that depends upon the colligative properties of the solution such as osmotic pressure measurements or electron microscopic counting of particles.

The number average molecular weight of a system,  $(M)_n$  is defined as:

$$(M)_n = (\sum M_i N_i) / \sum N_i$$

where  $M_i$  is the molecular weight of a given component  $I$  and  $N$  is the number of particles with molecular weight  $M$ . If we express  $(M)_n$  in Mole Fraction ( $X$ ) terms we get:

$$(M)_n = \sum M_i X_i$$

which in weight concentration terms ( $c = NM$ ) gives

$$(M)_n = (\sum c_i) / (\sum c_i / M_i)$$

### Weight Average Molecular Weights

A “Weight” average molecular weight is given by a technique where the experiment determines the “mass” of a particle, such as sedimentation, light scattering or gel electrophoresis where the staining is proportional to the number of amino acids in the particle [usually the case].

The weight average molecular weight of a system,  $(M)_w$  is defined as:

$$(M)_w = (\sum M_i c_i) / \sum c_i$$

which as a weight fraction,  $w$  becomes

$$(M)_w = \sum M_i w_i$$

or in Molar terms becomes:

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase

## Construction & Expression Module

$$(\overline{M})_w = (\sum M_i^2 N_i) / (\sum M_i N_i)$$

**For a pure protein of course  $(\overline{M})_n = (\overline{M})_w$**  and thus a comparison of molecular weight determinations using techniques that give number average and weight average molecular weights can be a criterion of purity.

### Basic Principles of Protein Structure

The essential dogma that needs to be considered is that the amino acid sequence of a protein in some way governs its three-dimensional structure and hence biological function. The three dimensional structure of a protein also governs its size and shape properties as well as its overall charge properties. The experimental techniques that are introduced in this laboratory are the ways that information about size, shape and charge are obtained for a protein. They are also techniques that can tell us about various levels of protein structure: primary structure, tertiary structure and quaternary structure..

In the first part of this section we will consider experiments that are based in some way on the charge on a protein, whether it is the native charge or some charge imposed by another molecule [usually an anionic detergent such as sodium dodecylsulfate, SDS] These include the techniques of Native Polyacrylamide gel electrophoresis [PAGE], and SDS-PAGE. The principle difference between these three electrophoretic techniques is that in the case of SDS-PAGE the native structure of the protein has been broken down by the SDS and a uniform charge per amount of protein introduced, and the separation is based solely on the size of the protein. In native PAGE separation is based on a complex relationship between charge, size and shape. In the last laboratory we considered Isoelectric focusing where the separation is based upon charge properties alone. At the end of this first section we will explore two-dimensional electrophoretic techniques that, for example, make use of separation on the basis of native charge properties and on denatured, size properties. Such techniques provide powerful ways to both analyze complex mixtures of proteins and to establish the purity of a given protein sample.

In the second part of this section, experiments will explore the relationship between native molecular weights and denatured molecular weights of proteins and how such comparisons establish the quaternary structure of a protein. Because under native conditions shape plays a role in the behaviour of the protein these techniques are in reality giving a parameter that is related to both shape [compact, globular, elongated etc] and size [molecular weight] and hence information about both shape and size can be obtained by using a combination of experiments.

### Molecular Weight Determination

#### *INTRODUCTION*

The determination of the molecular weight of a molecule would appear to be quite a straightforward procedure. For small molecules this is true: Elemental analysis and approaches using colligative properties -yield easily interpretable answers. The situation with a protein is somewhat different: If the primary sequence is known, it is a quite simple matter to determine the molecular weight from the amino acid composition if a single chain is present. Many proteins, however, contain subunits that may or may not be chemically similar. The molecular weight is then either a multiple of the molecular weight of the monomeric polypeptide chain or a combination of different polypeptide chains. Thus, even if the sequence is known (or sequences if more than one type exist) for the polypeptide chain, one does not necessarily know the molecular weight of the native molecule. Distinctions must be drawn between the native molecular weight and the denatured molecular weight. The

# **MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase**

## **Construction & Expression Module**

native molecular weight is that as the protein exists under normal in vivo conditions where it exhibits normal and full activity. The denatured molecular weight can be defined as that of the minimum covalently bonded structure that the molecule is broken down to under denaturing conditions such as the presence of SDS or guanidinium hydrochloride. This molecular weight must not be taken as the polypeptide chain molecular weight since two or more chains may be covalently linked by disulfide bridges and may contain covalently bound non-polypeptide entities such as carbohydrates, which contribute to the molecular weight. As we shall see, carbohydrate substituents can lead to anomalous molecular weight estimates using some of the approaches described. Similarly, most of the methods for determining molecular weights make some assumption concerning either the shape or the behavior of the protein relative to some reference: deviations in behavior from this reference lead to misleading molecular weight estimates. These estimates are, in many instances, averages, and depending on how the averaging is done, different answers for mixtures of species of different molecular weights are obtained. The various types of molecular weight averages that can be experimentally determined are described in detail here. Different techniques give different molecular weight averages, and a comparison of those obtained by approaches giving different types of molecular weights average also yield an idea of purity.

The concept of protein purity is somewhat nebulous: As sensitivity of protein detection has been increased, especially in techniques such as polyacrylamide gel electrophoresis (PAGE), it has also been far easier to detect impurities. Many of the experimental approaches for determining molecular weights are also used to estimate purity. Careful consideration of these methods is thus helpful in the context of examining purity as well as determining molecular weights.

SDS PAGE Protocol

Size Exclusion Chromatography Protocol

Native PAGE Protocol

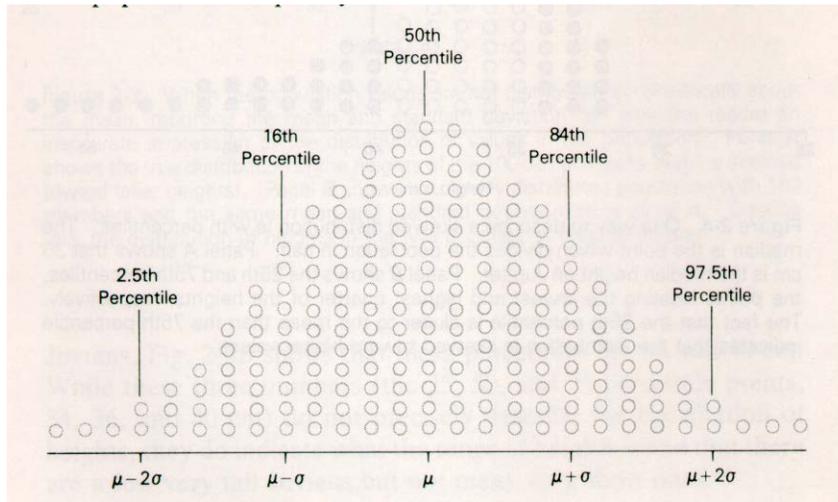
Western Blotting Protocol

### **Comments about Data Analysis: Averages and Standard Deviations**

Using Standard Deviations

Above we discussed putting the 95% confidence limits on a graph: where do these come from? The 95% confidence interval is the range of values that will describe 95% of all estimates of the measurement and is related to the standard deviation. Simply put 68% of the values will fall between the mean-one standard deviation and the mean+one standard deviation, while 95% of the values will fall between the mean-two standard deviations and the mean+two standard deviations. Thus using standard deviations not only tells you about how "good" the data is: a small standard deviation is "better" than a large standard deviation, but also allows the comparison of two numbers. This concept is illustrated graphically below for a "normal" distribution. Shown superimposed on this normal distribution are the so called percentile positions: the position where for example the bottom 2.5% of all estimates will fall.

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module



The mean value is the 50% percentile and shown are the positions of the 16% and the 84%, which define  $\pm$  one standard deviation and the 2.5% and 97.5%, which define the positions of  $\pm$  two standard deviations.

If you compare two normal distributions and take 95% confidence limits: ie the values laying between the 2.5% and 97.5% values, of each distribution and the values do not overlap, there is only 2.5% of the possible estimates of the higher number that could, in a normal distribution fall into the upper echelons of the normal distribution of the lower number: the numbers would be said to differ at the level of the 95% confidence limits, or at a p value of 0.025.

This approach can be used when comparing two sets of numbers, ie means and their standard deviations or two sets of parameters from the fit to an equation, for example the slope of a line and its standard deviation.

Graphs or table of numbers should include some discussion and indication of the statistics of the data shown.

## Obtaining Data for Use in Averages and Standard Deviation Calculations (and practicing pipetting!)

### Protocol:

Using a plastic weigh boat, pipet six aliquots of the chosen volume, taring the balance to zero before each addition. Record the weights of each addition The average weight of water pipetted for each of you chosen volumes can be calculated from the formula:

$$\bar{x} = \frac{\sum x_i}{n}$$

where  $\bar{x}$  is the mean, or arithmetic average,  $\sum x_i$  is the sum of the individual values, and  $n$  is the number of data values.

An estimate of the sample standard deviation,  $s$  of the values is given by:

$$s = \sqrt{\frac{\sum (x_i - \bar{x}_{av})^2}{(n-1)}}$$

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

where  $n-1$  is the number of degrees of freedom.

For example: consider that you obtained the following data for a micropipette set at 100 $\mu$ L:

Reading #	Weight, gm	Indicating Volume of:
1	0.102	0.102mL
2	0.099	0.099
3	0.103	0.103
4	0.098	0.098
5	0.101	0.101
6	0.101	0.101

The average,  $x_{av}$ , would be: 0.10067gm, indicating an average measurement of 100.67 $\mu$ L

To calculate the standard deviation:

Number	$x_i - x_{av}$	$(x_i - x_{av})^2$
1	1.33 $\mu$ L	1.769
2	1.67	2.789
3	2.33	5.429
4	2.67	7.129
5	0.33	0.109
6	0.33	0.109

Thus  $\Sigma(x_i - x_{av})^2 = 17.334$  and  $\Sigma(x_i - x_{av})^2 / (n-1) = 3.467$

Thus the sample standard deviation, which is the square root of  $\Sigma(x_i - x_{av})^2 / (n-1)$  is equal to 1.86

To obtain the scatter of a small data set, the estimate of the standard deviation is multiplied by the "Student's t" factor to give the confidence interval (CI) for a single value.

You should look up a table of student's t factors and determine the 95% confidence limits for your data.

## Worked Examples of Quantitative Analysis of the Data and Calculations Using Specific Activity

### Examples of Using Protein Concentration Determination in Conjunction with Activity Measurements:

#### Sample Data Set 1

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

The enzyme "Fucosyltransferase" catalyzes the transfer of Fucose from GDP-Fucose to a variety of acceptor substrates. An extract of parotid glands contains an active Fucosyltransferase.

## Experiment 1:

In the Bradford protein assay, dilutions of a 0.1mg/mL stock BSA (Bovine serum albumin) solution were made as follows.

$\mu\text{L}$ BSA	$\mu\text{L}$ Buffer	Measured OD
20	80	0.135
40	60	0.222
60	40	0.289
80	20	0.343
100	0	0.391

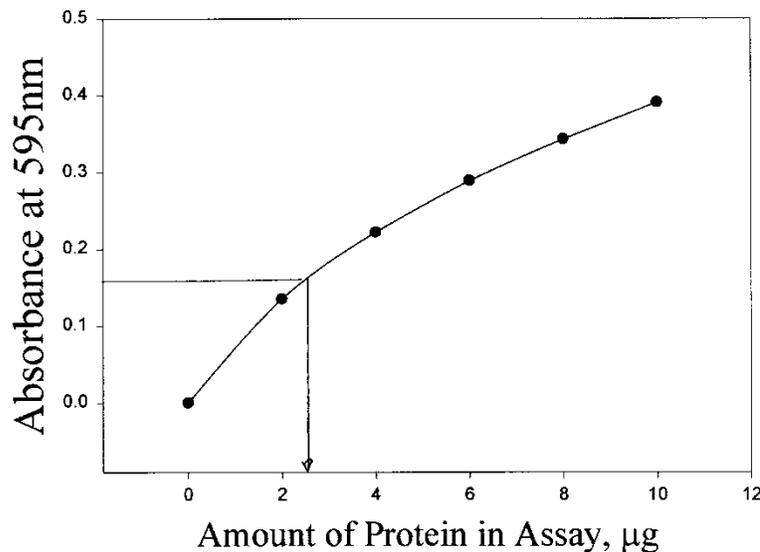
After incubation, in a total volume of 1mL, with Bradford reagent, the optical densities shown were recorded. For each measurement the spectrophotometer was "blanked" on a buffer blank.

In the first part of the problem you need to plot a standard curve for the Bradford Protein assay. Calculate the amounts of BSA added. You know that 1mL contains 100 $\mu\text{g}$ . Thus 20 $\mu\text{L}$  contains:  $(100 \times 20)/1000 = 2\mu\text{g}$  and so forth

Plot the graph as shown below. From the graph you can extrapolate any absorbance measurement to an amount of protein.

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

## Standard Curve for Bradford Assay



### Experiment 2:

10µL of a 5-fold dilution of a crude extract of parotid glands gave an optical density of 0.135. After purification, 100µL of the pure, undiluted preparation gave an absorbance of 0.153 under similar conditions.

An absorbance of 0.135 is = 2µg of protein. This was contained in 10µL or a 5-fold dilution. The original solution would have contained:

$$2 \times 100 \times 5 = 1000\mu\text{g of protein per mL}$$

The absorbance obtained from 100µL of the pure protein was 0.153. Extrapolation from the standard curve indicates that this is produced by 2.6µg of protein. The pure solution thus contains  $2.6 \times 10 = 26\mu\text{g/mL}$  of protein

### Experiment 3:

The enzyme is assayed by following the transfer of  $^{14}\text{C}$  fucose into the product:



The specific *radio-activity* of the GDP-Fucose is 125,500 cpm/mole.

## MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

In a time course experiment with pure enzyme, the following data was obtained for identical incubations containing 10 $\mu$ L of the stock "pure" enzyme, in a total of 100 $\mu$ L.

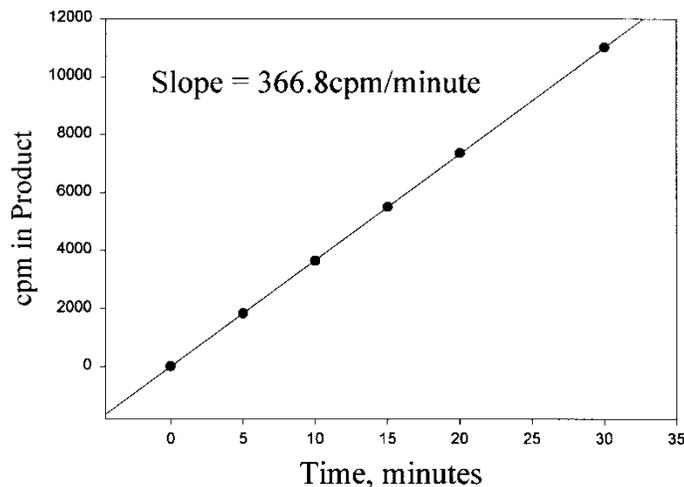
Time, minutes	cpm in Product
5	1820
10	3640
15	5500
20	7358
30	11002

10 $\mu$ L of the crude extract, assayed under identical conditions gives 2482 cpm in a 20 minute assay.

The plot of the time course of product production shows that it is linear with slope ( ie initial rate) of 366.8cpm/minute. This can be converted to moles of product using the specific radioactivity:  $366.8/125,500 = 2.92\text{mMoles/minute}$  in 100 $\mu$ L which would be a concentration of  $29.2\text{mMoles/mL/minute} = 29.2\text{M/minute}$

This rate was produced by 10 $\mu$ L of pure enzyme in 100 $\mu$ L, ie 100 $\mu$ L/mL and hence the pure enzyme concentration to give 29.2M/minute rate was  $(26 \times 100)/1000 = 2.6\mu\text{g/mL}$

### Rate Plot for Fucosyltransferase



The specific Activity of the pure enzyme is thus  $29.2/2.6 = 11.23\text{M/minute}/\mu\text{g/mL}$

You could have left the activity measurement in cpm/minute, in which case the specific activity would have been  $1410.8\text{cpm/minute/mL}/\mu\text{g/mL}$

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

Calculate:

- 1: The protein concentration of the pure stock solution of fucosyltransferase
- 2: The specific activity of the pure enzyme
- 3: The amount of fucosyltransferase per mL of the crude extract

10 $\mu$ L of the crude enzyme gave 124.1cpm/minute, 10 $\mu$ L of the pure enzyme gave 366.8cpm/minute: 10 $\mu$ L of the pure enzyme contains 0.26 $\mu$ g of protein. Thus the crude preparation would contain  $(0.26 \times 124.1)/366.8 = 0.088\mu$ g of fucosyltransferase per 10 $\mu$ L and hence 8.8 $\mu$ g of fucosyltransferase per mL.

and 4: The ?-fold purification necessary to obtain pure enzyme from the crude extract.

Since you know from above that the total protein concentration of the crude is 1000 $\mu$ g/mL and have just calculated that there are 8.8 $\mu$ g/mL of enzyme, the fold purification is simply  $1000/8.8 = 113.6$  fold. You could reach the same conclusion by calculating the specific activity of the crude enzyme and using the specific activity of the pure as well.

## Sample Data Set 2

Saliva is a fluid secreted by two major types of gland in the mouth, the sublingual [SL] glands and the submaxillary [SM] glands. Secretions from each gland were collected, at a rate of 1.3mL per minute for SL, and a rate of 2.7mL per minute for SM secretions, and the activity of the enzyme amylase measured by determining the amount of p-nitrophenol(pNP) released from pNP-Glucoside, using absorbance measurements at 410nm, as a function of time with 50 $\mu$ L additions of the saliva samples, in a total volume of 1mL.

The following data was obtained:

Time, minutes	OD [SM glands]	OD [SL glands]
0	0	0
0.5	0.093	0.062
1.0	0.183	0.125
1.5	0.279	0.183
2.0	0.354	0.252
3.0	0.517	0.376
5.0	0.722	0.629

Using the BCA Protein assay which gives a color at 562 as the result of a chemical reaction, the following data was obtained for the standard curve:

$\mu$ L BSA, 1mg/mL Stock	OD @ 562nm
0	0.052

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase

## Construction & Expression Module

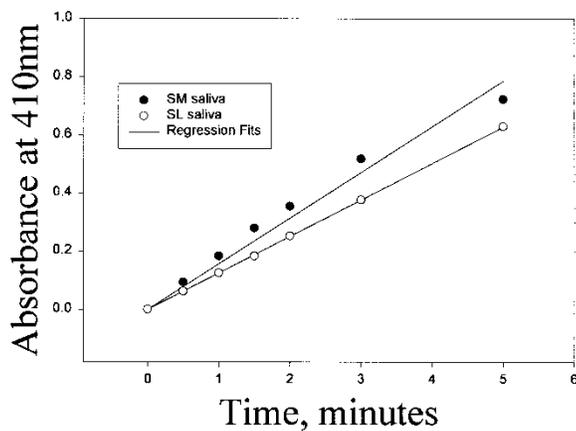
10	0.145
20	0.238
30	0.331
40	0.424
50	0.518
75	0.752
100	0.983

In a separate experiment, using identical conditions, 100 $\mu$ L of the SL and the SM extract gave OD values of 0.191 and 0.784 respectively.

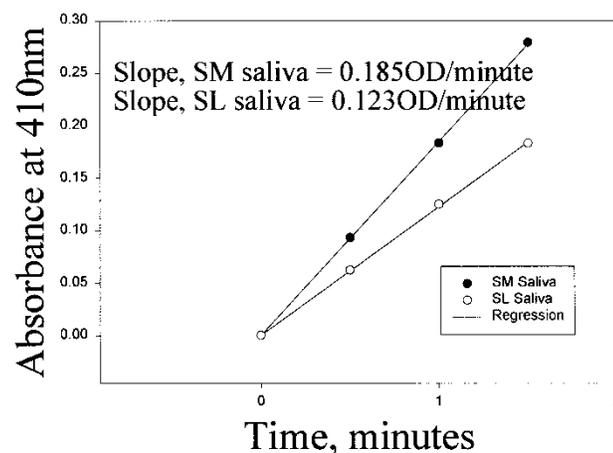
a) Which secretion contributed the most amylase activity per mL of total saliva?

First plot the product versus time plots to determine whether the plots are linear and you have initial rates:

### Rate Plots for Glucosidase Activity



### Rate Plot for Glucosidase Truncated Data to Obtain Linearity



# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

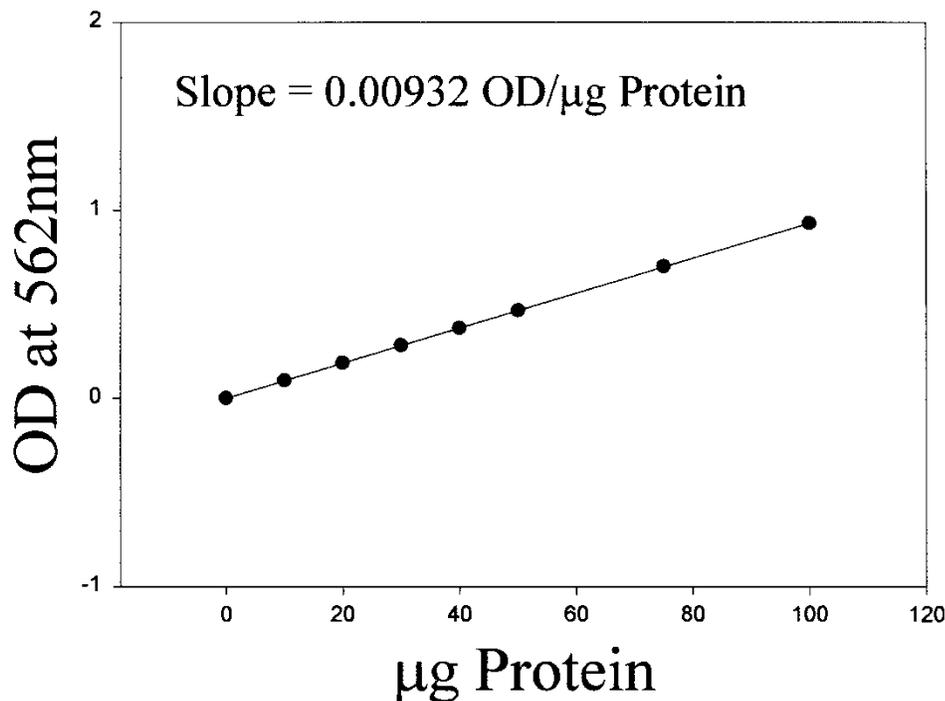
Second, it is clear that the SM saliva is the more active per 50 $\mu$ L added to the assay and hence contributes most to the total saliva since it is also secreted faster.

and b) what was the specific activity of the amylase in each secretion?

From the above plots the rates for each saliva are 0.185OD/minute and 0.123OD/minute for SM and SL saliva respectively when 50 $\mu$ L aliquots are added

From the BCA standard Plot shown below the slope is 0.00932OD/ $\mu$ g and hence from the measured Ods under identical conditions it is clear that 100 $\mu$ L of the secretions contain  $(0.191-0.052)/0.00932 = 14.9\mu\text{g}$ , and  $(0.784-0.052)/0.00932 = 78.5\mu\text{g}$  respectively for SL and SM secretions. Thus 50 $\mu$ L would contain half these amounts: : 7.45 and 39.3 $\mu$ g respectively. Since both rate and protein are now calculated on a basis of 50 $\mu$ L, simply divide rate by protein to obtain specific activity:

## Plot of BCA Standard Curve



SM: Specific Activity =  $0.185/39.3 = 0.0047\text{OD/minute}/\mu\text{g Protein}$

SL: Specific Activity =  $0.123/7.45 = 0.0165\text{OD/minute}/\mu\text{g Protein}$

The millimolar extinction coefficient of pNP is  $18\text{cm}^{-1}$

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

You can also convert OD/minute to mM pNP produced/minute by dividing by 18

## Sample Data Set 3

A particular receptor protein is defined by its ability to bind to radiolabeled GTP.

You are trying to determine the concentration of this protein in an unknown sample. You have a stock solution of BSA at 1.1mg/mL and a bottle of "Bradford" reagent. You set up a standard curve with 100 $\mu$ L of a series of 2-fold dilutions of your standard protein, including in the curve 100 $\mu$ L of the undiluted protein, add dye, incubate and measure the absorbance at 595nm. You obtain the following data:

Serial Dilution	Absorbance at 595nm
None	1.23
Dilution 1	0.87
Dilution 2	0.51
Dilution 3	0.27
Dilution 4	0.15
Dilution 5	0.08
Dilution 6	0.04
Dilution 7	0.02
Dilution 8	0.01
Blank: No Protein	0

In three different determinations using 20 $\mu$ L, 50 $\mu$ L and 100 $\mu$ L of the unknown you obtain the following data:

$\mu$ L Unknown	Absorbance at 595nm
20	0.49
50	0.96
100	1.48

## What is the concentration of the unknown protein in mg/mL?

200 $\mu$ L of this protein in a dialysis bag was equilibrated with radiolabeled GTP [specific radio-activity of 2,000 cpm/nmole] at a concentration sufficient to saturate the protein with GTP. When the contents of the bag and the dialysis buffer were analyzed the following results were obtained:

200 $\mu$ L Bag contents: 125,269cpm  
200 $\mu$ L Dialysis Buffer: 93,847cpm

What is the Specific Activity of the protein in terms of moles GTP bound/mg protein?

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

## A Worked Example of Using Specific Activities

### Using Activity Measurements to Determine the Amount of Protein in a solution

Experiment 2:

10  $\mu$ L of a 5-fold dilution of a crude extract of parotid glands gave an optical density of 0.135. After purification, 100  $\mu$ L of the pure, undiluted preparation gave an absorbance of 0.153 under similar conditions.

An absorbance of 0.135 is = 2  $\mu$ g of protein. This was contained in 10  $\mu$ L or a 5-fold dilution. The original solution would have contained:

$$2 \times 100 \times 5 = 1000 \mu\text{g of protein per mL}$$

The absorbance obtained from 100  $\mu$ L of the pure protein was 0.153. Extrapolation from the standard curve indicates that this is produced by 2.6  $\mu$ g of protein. The pure solution thus contains  $2.6 \times 10 = 26 \mu\text{g/mL}$  of protein

Experiment 3:

The enzyme is assayed by following the transfer of  $^{14}\text{C}$  fucose into the product:



The specific radio-activity of the GDP-Fucose is 125,500 cpm/mole.

In a time course experiment with pure enzyme, the following data was obtained for identical incubations containing 10  $\mu$ L of the stock "pure" enzyme, in a total of 100  $\mu$ L.

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase Construction & Expression Module

Time, minutes	cpm in Product
5	1820
10	3640
15	5500
20	7358
30	11002

100  $\mu$ L of the crude extract, assayed under identical conditions gives 2482 cpm in a 20 minute assay.

The plot of the time course of product production shows that it is linear with slope ( ie initial rate) of 366.8cpm/minute. This can be converted to moles of product using the specific radioactivity:  $366.8/125,500 = 2.92\text{mMoles/minute}$  in 100  $\mu$ L which would be a concentration of  $29.2\text{mMoles/mL/minute} = 29.2\text{M/minute}$

This rate was produced by 10  $\mu$ L of pure enzyme in 100  $\mu$ L, ie 100  $\mu$ L/mL and hence the pure enzyme concentration to give 29.2M/minute rate was  $(26 \times 100)/1000 = 2.6\mu\text{g/mL}$

## References

Robert L. Switzer, Liam F. Garrity, 1999, "Experimental Biochemistry" page 94

William W. Wilfinger, Karol Mackey, and Piotr Chomczynski, Effect of pH and Ionic Strength on the Spectrophotometric Assessment of Nucleic Acid Purity: *BioTechniques* 22:474-481 (March 1997)

J.R. Whitaker and P.E. Granum, *Analytical Biochemistry* 109, 156-159(1980)

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. 1995. How to measure and predict the molar absorption coefficient of a protein. *Prot. Sci.* 4:2411-2423.

Pace, C.N. and Schmid, F.X. 1997. How to determine the molar absorbance coefficient of a protein. *In Protein Structure: A Practical Approach.* (T.E. Creighton, ed.) pp. 253-259. IRL Press, Oxford, U.K.

[Nozaki Y.](#) "Determination of the concentration of protein by dry weight--a comparison with spectrophotometric methods" . [Arch Biochem Biophys.](#) 1986 Sep;249(2):437-46.

# **MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase**

## Construction & Expression Module

Thermo Scientific Pierce Protein Assay Technical Handbook:  
[http://www.piercenet.com/files/1602063\\_PAssayHB\\_122910.pdf](http://www.piercenet.com/files/1602063_PAssayHB_122910.pdf)

# MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase

## Construction & Expression Module

<i>Learning Objective</i>	<i>Excellent</i>	<i>Accept</i>	<i>Poor</i>	<i>Fail</i>
<p><b>Can calculate yields and amounts of materials given appropriate information:</b>                      eg: The turnover number of pure MDH (Mol Wt 34,400) is 160000. What is the concentration of MDH in a solution if 10<math>\mu</math>L of a 1000 fold dilution, added to 3mL of assay mix gave a <math>\Delta A_{340}</math> of 0.743/min. <b>5mL of this stock was purified from 40mL of crude extract. 20<math>\mu</math>L of a 100 fold dilution gave <math>\Delta A</math> of 0.799/min. What is the yield of the purification?</b></p>				
<p><b>Can accurately use serial dilutions and dilution factors to calculate concentrations and amounts of materials in appropriate solutions</b>                      eg: Two 10fold serial dilutions of an NADH solution gave an absorbance of 0.042 when 100<math>\mu</math>L was diluted into a total of 1mL. If 48.7mg of NADH had been used to make 10mL of this solution, what was the purity of the NADH</p>				
<p><b>Calculate averages and standard deviations and errors and appreciate the difference between them</b></p>	Understands the difference between Population and Sample standard Deviations and the relationship between SD (scatter of the data) and SE (preciseness of population mean). Understands that Variance is the ave of $\Delta^2$ from the mean			
<p><b>Understands the concept of a normal distribution</b></p>	The curve is symmetric at the center (i.e. around the mean, $\mu$ ). Exactly half of the values are to the left of center and exactly half the values are to the right, and the mean, mode and median are all equal			
<p><b>Understands Confidence limits</b></p>	Uses the fact that for a normal distribution, 68% of the data falls within 1 SD of the <u>mean</u> , .95% within 2 SD & 99.7% of the data falls within 3 SD. Uses to indicate the confidence level in the difference between 2 numbers			
<p><b>Can discuss and use, as appropriate, 1 and 2 way ANOVAs</b></p>	Appreciates that 1 way ANOVA is used when comparing 1 independent variable with multiple conditions and 2 way applies to 2 independent variables and multiple conditions			
<p><b>Understands that an experiment, no matter how “good” the data, needs to be repeated</b></p>				

The molecular weight of NADH is 665.4, the mM Extinction coefficient at 340nm is 6.22

# **MCC: Exploring Structure-Function Relationships in Malate Dehydrogenase**

## Construction & Expression Module