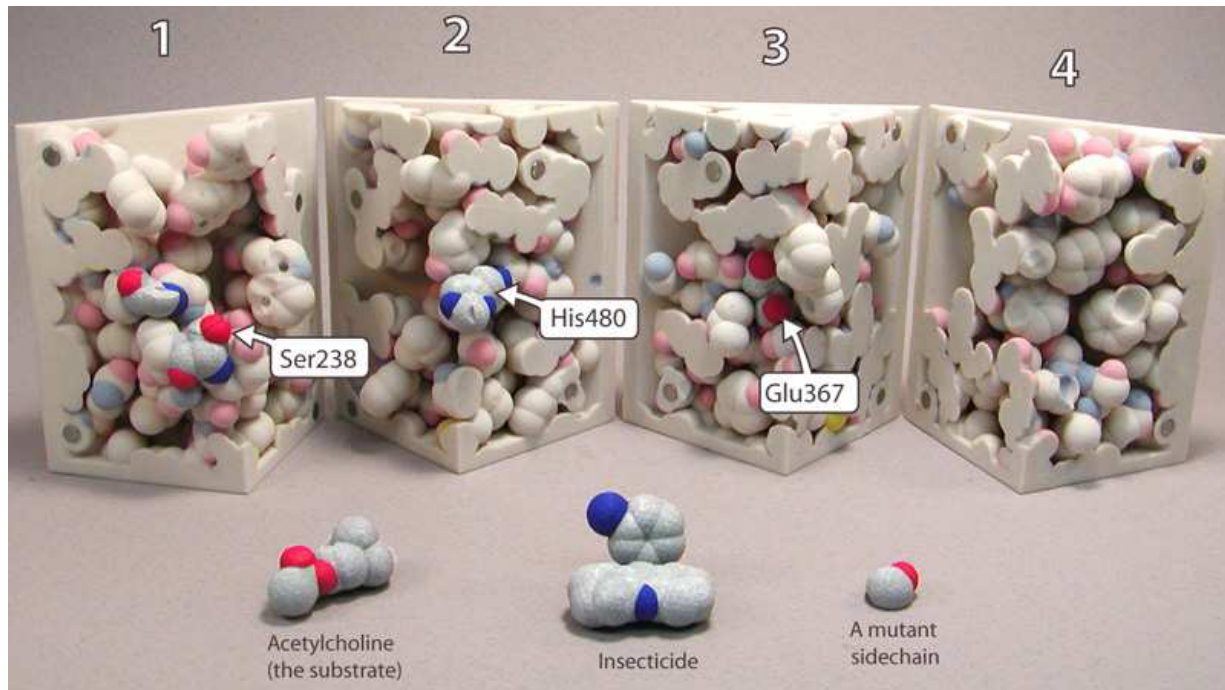


Evolution in Action: Modeling Insecticide Resistance in Mosquitoes



Instructor's Manual

Classroom materials were developed by Margaret Franzen, Pellissippi State Technical Community College, Knoxville TN and Center for BioMolecular Modeling, Milwaukee School of Engineering, Milwaukee WI, in collaboration with Lynda Jones, Caitlin Gabel School, Portland OR. Jmol files were developed by David Goodsell of The Scripps Research Institute, La Jolla, CA. Physical models were developed in collaboration with Tim Herman, Center for BioMolecular Modeling (CBM), Milwaukee School of Engineering, Milwaukee WI. Funding for this project was provided in part by NSF CCLI grant #0442409 and NIH NCRR SEPA award RR022749.

Permission is given to modify and/or copy these materials for classroom use, but they may not be sold or published. Feedback on how the materials work in your classroom would be greatly appreciated. Additional materials are being developed; feel free to contact us for the latest updates.

Correspondence should be sent to Margaret Franzen, Center for BioMolecular Modeling, Milwaukee School of Engineering, Milwaukee WI 53202, franzen@msoe.edu.

Concepts Addressed by Acetylcholinesterase Module

- The active site of a protein is not necessarily formed by adjacent amino acids in the peptide chain
- Protein structures are not static, but flexible
- Mutation occurs prior to natural selection
- Some mutations at the DNA level are not expressed at the protein level (silent mutations)
- Not all mutations are 'bad'
- Genes can have alternate introns splice sites, leading to more than one protein from a single gene

Overview of Materials and Teaching Points

These materials explore the structure-function relationship of a single protein, acetylcholinesterase. The acetylcholinesterase gene and the protein it encodes can be used to demonstrate a number of biological concepts, including enzyme specificity, competitive inhibition, mutation, characteristics of the genetic code, alternate splice sites, natural selection, bioinformatics, and disease transmission. Many of these materials can be presented either at an introductory or an advanced level. For instance, materials are provided to allow students to compare DNA and protein sequences by hand, or these comparisons can be completed using programs available on the Internet. Furthermore, it is not necessary to incorporate ALL the materials in the classroom at one time. We often use acetylcholinesterase as one of a few 'anchor' proteins that serve as examples to help tie together a number of biological concepts. My hope is that when students leave my class, they can provide specific examples of a few proteins that illustrate a number of biological concepts.

Nerves send signals to muscles by way of neurotransmitters, such as acetylcholine. Once the signal has been received, the neurotransmitter must be broken down and recycled, or the cell would not be able to distinguish a new signal as it enters the receptors. Acetylcholinesterase, which breaks down acetylcholine into acetic acid and choline, is the target of nerve agents, such as sarin, as well as insecticides, including DDT and Malathion. For more background information on acetylcholinesterase, check out David Goodsell's Molecule of the Month on the Protein Databank website: http://www.rcsb.org/pdb/education_discussion/molecule_of_the_month/download/Acetylcholinesterase.pdf.

Aside from being pesky nuisances, mosquitoes often transmit blood-borne diseases, such as malaria and West Nile virus. In an attempt to control the spread of disease, insecticides have been used widely. The acetylcholinesterase gene has mutated in some mosquitoes, resulting in resistance to insecticides. When mosquitoes are exposed repeatedly to insecticides, those with resistance to the chemicals survive to reproduce, resulting in a population of mosquitoes that are no longer able to be controlled with the insecticide. This material explores the nature of the mutation leading to resistance at the molecular level and is based on a paper published by Weill, et al. (2003).

First, students explore the structure of acetylcholinesterase, utilizing computer simulations (located at <http://www.scripps.edu/mb/goodsell/jmol/ach>), as well as physical models, available from the MSOE Model Lending Library. They discover that the catalytic triad consists of serine, histidine and glutamic acid residues that are separated by many amino acids in the single polypeptide, but that the three residues lie physically near each other, buried at the base of a long, narrow pocket within the protein. Although the substrate fits neatly in the active site, physically in contact with each of the residues in the catalytic triad, students quickly discover that the only way for the substrate to access the active site is if the protein undergoes a conformational change – appearing to 'breathe' to allow the substrate to reach the catalytic site.

After exploring the active site of acetylcholinesterase, students compare the DNA sequence from two strains of mosquito, one sensitive, and the other resistant, to insecticides. They discover that there are 28 nucleotide differences between the two strains. Next, they align the protein sequences, and discover that there is only ONE difference between the two strains. This demonstrates silent mutation – changes at the DNA level that are not expressed as changes in the protein level. This is a great opportunity to review the characteristics of the genetic code – it is a degenerate code - more than one codon can code for a specific amino acid), and that there is third base 'wobble' – often mutations in the third base position still code for the same amino acid.

After examining acetylcholinesterase bioinformatics, students then go on to study the impact of the specific amino acid change on the protein function, returning to the computer simulations and physical models. They discover that the mutation is not in the catalytic site of the protein, but in the long narrow pocket leading to the catalytic site. Insecticides serve as competitive inhibitors, blocking the path to the active site of the molecule. Students discover that the mutation is from a glycine residue to a serine residue, and that this minor change still allows the substrate to squeeze through the gorge. They also discover that, although the residue in the resistant strain is only slightly larger, it is sufficiently large to block the inhibitor from entering the gorge.

After students have discovered the nature of the mutation and the impact on protein structure, you could have them go one step further by asking them to propose a design for a better insecticide – one that would inhibit the enzyme regardless of what mutations might arise. Students may suggest an inhibitor that binds directly at the catalytic site, providing another teaching opportunity. (Insects with mutations at the catalytic site would not survive, since the protein would not function normally.)

This manual is organized in sections for those instructors interested in only presenting a portion of the material. Most sections are divided into subsections. Throughout the material, Section A is more appropriate for introductory courses. Students are provided with the aligned sequences and are required to identify the differences between the sequences. In

section B, students actually utilize Internet tools for bioinformatics to align the sequences. In some cases, there are additional suggestions in section C to allow for more in depth study of bioinformatics.

I Background on Acetylcholinesterase

It is strongly recommended that students begin by learning a little about the function of acetylcholinesterase. A great place to start is at the Molecule of the Month website sponsored by the Protein Data Bank. Acetylcholinesterase information is found at: http://www.rcsb.org/pdb/static.do?p=education_discussion/molecule_of_the_month/pdb54_1.html.

- A. Utilizing computer visualizations and physical models
 1. You can access computer visualizations of acetylcholinesterase structure for both sensitive and resistant strains at: <http://www.scripps.edu/mb/goodsell/jmol/ach/>. There are three different links on this page, described below. At each of these links, there are JMol images that allow the viewer to rotate the structure, change the appearance (backbone, space fill, etc.), and zoom in on important features. (No additional software is required to view JMol files, which is a distinct advantage over Chime files.) If you click on an atom in the structure, its identity will appear in the footer at the lower left of the browser window.
 - a) 'Basics' explores the structure of acetylcholinesterase, showing the long narrow gorge, or pocket, leading to the active site, which is buried deep within the protein. Acetylcholine is in green and is positioned in the active site.
 - b) 'Active site' explores in greater detail, the active site of the protein, including the three amino acids that form the catalytic triad (SER200, GLU327 and HIS440)¹. Additional amino acids that define the boundaries of the gorge are shown in grey and labeled 'other active site residues'. There are two images, one with the substrate, acetylcholine, in green, and the other with the inhibitor, colored yellow.
 - c) 'Mutant' focuses on comparing the active site of the wild type (sensitive to insecticide) and the mutant (resistant to insecticide). [This part of the visualization is explored in section VI, after students have determined the nature of the mutation rendering resistance to insecticides.]
 2. After exploring the computer models, students can explore the structure in three dimensional space, using the acetylcholinesterase model available from the MSOE Model Lending Library. These models can be reserved for use for 10-14 days (including shipping time); the borrower only pays for return shipping. Additional information about the lending library can be found at: <http://www.rpc.msos.edu/lib/>. It is possible to show students the computer visualizations alone, but our experience has been that students really don't understand fully until they manipulate the molecules themselves. These models serve as thinking tools, allowing students to ask much more meaningful questions about the structure-function relationship in this and other proteins.

II Exploring Acetylcholinesterase Gene Structure – Alternate Splice Sites

- A. Students can read the handout 'Acetylcholinesterase NCBI' on the CD to discover that there are two different forms of the enzyme, but only one gene. They will learn that which forms are expressed in various tissues, and they can see the map of the two splice variants, one which splices exon 4 to exon 5 (E4-E5) and the other that splices exon 4 to exon 6 (E4-E6).
- B. Students can access the information (and much more) themselves by going to the NCBI website (www.ncbi.nlm.nih.gov) and searching in the Gene section for acetylcholinesterase. If they locate the information on the acetylcholinesterase (Yt blood group) from *Homo sapiens*, they will find the information on the handout in the Summary as well as the section 'Genomic Regions, Transcripts and Products'. If there is time, students can dig much deeper by following links in NCBI. Depending on where they explore, they might discover that the two forms can be expressed in one tissue at different times (for instance, under stress), or that the acetylcholinesterase in the blood serves the purpose of binding organophosphates as well as any acetylcholine that was not removed at the synapse.

III Bioinformatics Part 1: DNA Alignment

- A. Using two pre-aligned sequences (one sensitive, one resistant to insecticide)

¹ These residue numbers are for acetylcholinesterase from *Torpedo californica* (electric ray) 1AMN. Computer simulations involving docked protein are from *Drosophila melanogaster* (fruit fly) 1QON, with catalytic residues SER238, GLU367 and HIS480, The corresponding catalytic residues in *Culex pipiens* (mosquito) are SER327, GLU453 and HIS567. Although there are some differences in the overall amino acid sequence in the three organisms, they maintain the same overall structure at the active site, allowing for meaningful comparison of structures. When the models were developed, the crystal structure for mosquito acetylcholinesterase had not yet been resolved.

1. See "Worksheets: SLA-B SR DNA Alignment" on the CD. Students use highlighters to denote differences in the two sequences.
- B. Using Bioserver to align two sequences (one sensitive, one resistant to insecticide)
 1. Go to the Bioserver website at <http://www.bioservers.org/bioserver/>.
 - a) Under 'Sequence Servers' click 'Register' (unless you have already registered at the DNA Learning Center and have a username and password). Registration is free, and it allows you to save your sequences for use in further comparisons.
 - b) Once you have registered, you may enter the Sequence Server.
 2. Place the acetylcholinesterase sequences in your work area:
 - a) Click on 'Manage Groups'.
 - b) Select "Public" under 'Sequence Sources'
 - c) Search for "Resistant Ache" and "Sensitive Ache" posted in 8/05 and click the box to the left of each of these listings.
 - d) Click 'OK'. This will save the sequences to your work area.
 3. After saving both sequences, align them using Clustal W:
 - a) Click on the box to the left of the sequences you want to compare (SLA-B and SR).
 - b) Select 'Align:ClustalW' from the drop down box near the top of the page, then click 'Compare'.
 - c) A new window will open with the alignment. Yellow boxes indicate positions in which the nucleotide sequences are different. You may print the comparison, but some printers don't pick up the yellow highlighting. After printing, you may need to use a highlighter to mark the differences between the sequences. Alternately, if you have a color printer available, you may cut and paste the alignment into a word processing document (the color highlights transfer), then print the transferred document.

IV Bioinformatics Part 2: Protein Alignment

- A. Aligning sequences from one sensitive and one resistant strain
 1. Students compare the aligned SLA-B and SR protein sequences and discover that, in spite of 28 differences in the DNA, there is only one difference in the protein. A "*" under the two aligned sequences indicates the amino acids are identical. This handout is called "Alignment of SLA-B and SR protein sequences"
- B. Translating the sequences, then aligning the translated sequences
 1. Translating sequences
 - a) In BioServer website (<http://www.bioservers.org/bioserver/>), select SLA-B and click the 'Open' button to the right of the sequence. Cut and paste this sequence into a word processing file.
 - b) Go to <http://bio.lundberg.gu.se/edu/translat.html>
 - c) Cut and paste the sequence into the box, and click translate. The protein sequence will appear in a new window. Next there will be the DNA sequence, with three rows of protein sequences below the DNA sequence. These protein sequences are staggered and correspond to the three reading frames. The one letter abbreviation for the amino acid is printed directly underneath the first nucleotide in the codon that codes for that amino acid.
 - d) Cut and paste the protein sequence (at the top of the output) into the word processing document below the DNA sequence. Be sure to label the source of the DNA.
 - e) Repeat steps a-d for the SR sequence.
 2. Aligning translated sequences
 - a) Go to www.expasy.ch/tools/sim-prot.html.
 - b) Under "Sequence 1" select 'User-entered sequence' and enter 'SLA-B' for 'Sequence Name'.
 - c) Cut and paste SLA-B DNA protein sequence (saved in the word processing file) into the 'paste your sequence here' box.
 - d) Under "Sequence 2" select 'User-entered sequence' and enter 'SR' for 'Sequence Name'.
 - e) Cut and paste SR DNA protein sequence (saved in the word processing file) into the 'paste your sequence here' box.
 - f) Click 'submit'.
 - g) The results will show the two aligned sequences. The symbol "*" below the aligned sequences shows identical amino acids in that position.

- V Bioinformatics Part 3: Comparison of Multiple Strains (This entire section is more advanced and can be omitted; this section just solidifies the finding that a single amino acid change results in insecticide resistance. If you have time to include this material, it provides very powerful evidence that students are on the right track!)
- A. Comparing pre-aligned DNA sequences from 30 strains (19 sensitive, 11 resistant to insecticide)
 - 1. Students search 30 aligned DNA sequences to locate differences. The SLA-B strain is used as the standard for comparison. Note that most of these sequences only contain exon 3; SLA-B and SR contain the complete gene sequence. Only exon 3 was amplified in the other strains once it was determined that the only amino acid difference between SLA-B and SR was in exon 3.
 - 2. After all differences are marked with a highlighter, students should locate any differences that are found in ALL of the resistant strains and NONE of the sensitive strains. They will identify one difference at base #739. Students can calculate the amino acid by dividing the nucleotide number by 3 (3 nucleotides per codon), and rounding the value to the next whole number if it is not evenly divisible by 3. They will discover that the amino acid (#247) is the ONLY amino acid difference identified between SLA-B and SR.
 - B. Aligning the 30 strains using Bioserver
 - 1. After logging into the Bioserver website (<http://www.bioservers.org/bioserver/>), students can select additional acetylcholinesterase sequences of strains that are sensitive and resistant for alignment by opening the drop down box(es) that say 'none' and selecting additional sequences. Each time a new sequence is opened in the list, a new drop down box appears.
 - 2. There are 29 strains sensitive, and 11 strains resistant, to inhibitors. The BioServer only allows eight sequences to be aligned at once, so must complete the alignment in segments, or instructors may assign the alignment of only some sequences. It is best to always include both SLA-B and SR in all alignments.
 - 3. After sequences are selected, click 'Compare" Align Clustal W. This will align the selected sequences. Note: if you print directly from the output, the color highlights typically don't transfer. You may avoid this problem by cutting and pasting the results into a word processing document, then printing the new document.
- VI Bioinformatics Part 4: Understanding the Genetic Code. This section could easily be omitted; it focuses on the location of each of the mutations within the codon and develops an understanding of both the degeneracy of the code and the wobble hypothesis.
- A. Using a prepared chart to identify the location of the mutations within the codon
 - 1. The file 'analysis of location of mutations within codons' provides a table listing all 28 nucleotide differences in the acetylcholinesterase gene between SLA-B and SR strains, listing both the nucleotide number and the codon containing the change. (The codons for SLA-B are listed.) The variant nucleotide is printed in red. The single change, resulting a different amino acid, is highlighted in yellow.
 - 2. This table can be used to discuss the location of each mutation in the codon (most are in the third base). Students can compare the codons and the amino acids for which they code. This is a perfect opportunity to reinforce the concept of the redundancy of the genetic code. Most of the mutations occur in the third codon; two of the three codons for which the mutation is in the first base, also code for the same amino acid.
 - B. Preparing their own chart to identify the location of the mutations within the codon
 - 1. Students can use the data they collected earlier to generate a list of the nucleotide differences and the codons in which each mutation occurs. They will need BOTH the aligned SLA-B and SR DNA sequences, as well as the translated sequences with the amino acids list.
 - 2. Using the DNA alignment output generated in section II.B.3, highlight the differences in the DNA sequences in resistant and sensitive strands on the translation output sheets generated in section III.B.1.
 - 3. The top reading frame of the protein sequences is the actual protein sequence. This particular translation program prints the one letter amino acid code directly under the middle base of the codon. Students can thus identify the codon that includes each of the mutations, as well as the amino acid for which it codes. This information can be tabulated; a copy of a suitable data table is found in 'data table VB3'.

VII Protein Structure: Determining the Structure-Function Relationship of the Mutation

- A. Utilizing computer visualizations and physical models
 - 1. Students can now access the computer visualizations again at <http://www.scripps.edu/mb/goodsell/jmol/ach/>. They may wish to review the 'basics' and 'active

- site' links before going on to the 'mutant' section. 'Mutant' focuses on comparing the active site of the wild type (sensitive to insecticide) and the mutant (resistant to insecticide). Both substrate (acetylcholine) and inhibitor are observed in the active site gorge. Students should discover that although both molecules fit in the active site in the wild type protein, the inhibitor is too large to fit in the pocket with the larger serine residue in the mutant protein.
2. After exploring the computer visualizations, students can return to the physical models to study the impact of the mutation on the ability of the substrate and inhibitor to fit in the active site gorge. They will quickly discover that the substrate is small enough to fit in the active site whether the normal or mutant residue is present, but that the inhibitor is too large to fit in the pocket in the presence of the mutation.
- B. Proposing a new insecticide structure
1. After students have worked with the computer visualizations and physical models, it is very easy to transition into applying their knowledge. Students can be asked to describe characteristics of an inhibitor that will still function in the presence of the mutation. There are several avenues they may explore – including, but not limited to, a smaller inhibitor, an inhibitor that actually binds at the active site (which CAN'T mutate and still be functional), or perhaps an inhibitor that binds at the top of the pocket, blocking access to the active site.

Most of these teaching points have been developed into a Waksman Challenge for high school students, which was offered in spring 2006. The link to the Waksman Challenge website is: <http://wakschallenge.rutgers.edu/index.php>. The acetylcholinesterase challenge will eventually appear under the 'archives'. (There is a copy of the challenge on the CD, in the folder 'Waksman Challenge' under 'Insecticide Resistance'.)

Bibliography

Weill, Mylene, Georges Lutfalla, Knud Morgensen, Fabrice Chandre, Arnaud Berthomieu, Claire Berticat, Nicole Pasteur, Alexandre Philips, Philippe Fort and Michel Raymond. 2003. Insecticide Resistance in Mosquito Vectors. *Nature* 423: 136-137. Available free online at: http://www-eve.ucdavis.edu/eve102/pdf/Weill_etal_Nature_2003.pdf