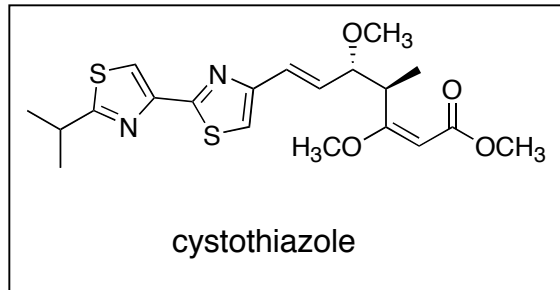


5.451 F2005
Problem Set 1.

1. (25 pts) Shown below is the structure of the cyclic peptide cystothiazole, a potent antibiotic, produced by the bacteria *Cystobacter fuscus*. Draw a reasonable biosynthetic pathway that includes chemical structures of intermediates and the names or abbreviation of the types of enzymes involved.

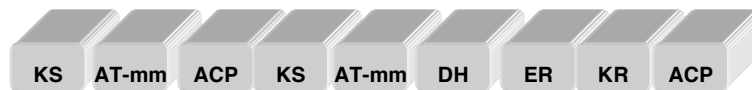
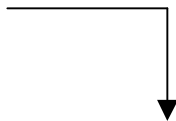


5.451 F2005
Problem Set 1.

2. (25 pts) Predict the structure of the polyketide that is encoded by the following set of enzymes.



enzyme 1



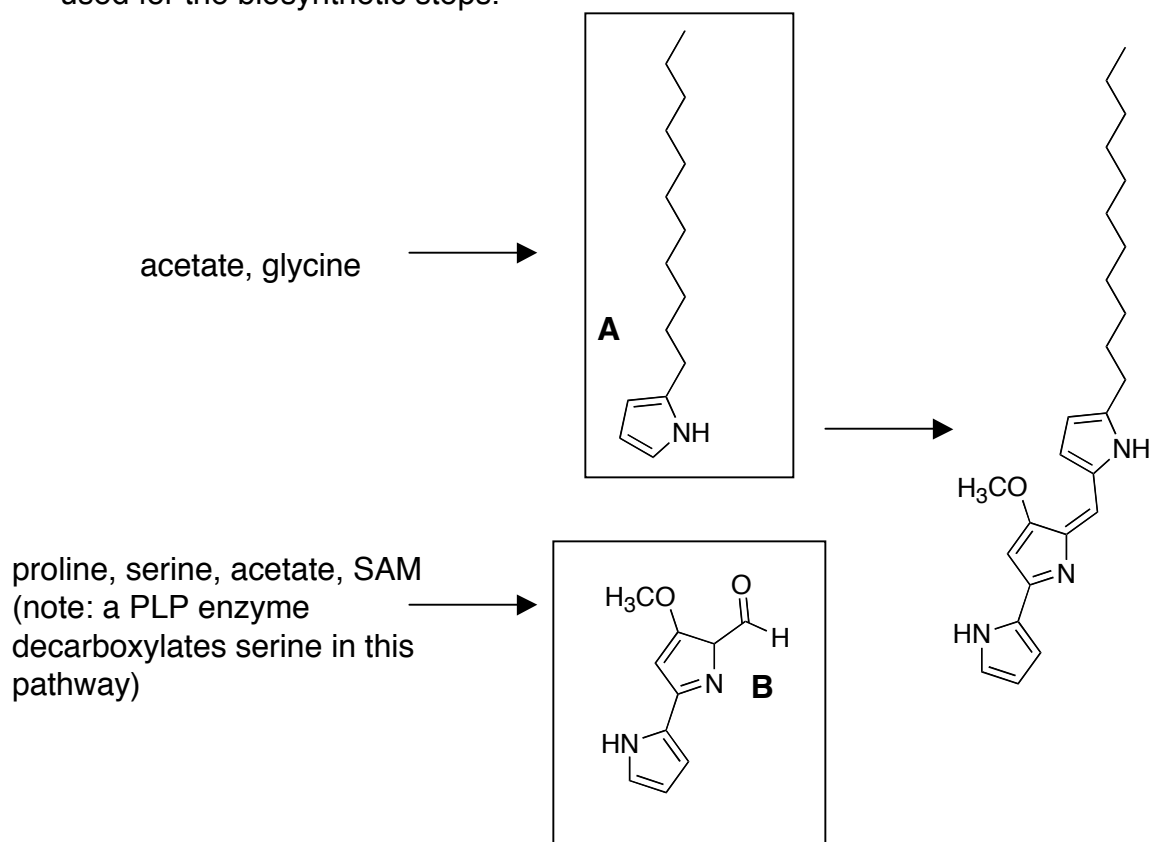
enzyme 2



enzyme 3

5.451 F2005
Problem Set 1.

3. (20 pts) Labeling studies have shown that the following molecule from *Streptomyces coelicolor* is made up of one unit of proline, one unit of serine, one unit of glycine and multiple units of acetate (i.e. malonyl CoA). The biosynthetic pathway is convergent, with the two intermediates A and B condensing to form the final product. Propose biosynthetic pathways for A and B and indicate the basic type of enzymes that would be used for the biosynthetic steps.



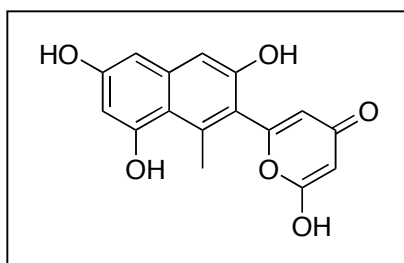
5.451 F2005
Problem Set 1.

4. (20 pts) Draw the thioester-tethered linear polyketide intermediate for the following compound, and show the steps that convert the intermediate into the final product. (Enzymes not required to be mentioned in the answer).

If the producing organism *Streptomyces glaucescens* is cultured in the presence of [1,2-¹³C₂] acetate, show the ¹³C labeling pattern expected to be observed in the final product using a bold line between the desired two carbons.



example



5.451 F2005
Problem Set 1.

5. (10 pts) Sequence homology- looking at the amino acid sequence similarity among proteins- is one modern method to predict the function of a new protein. How are homologues picked out of a database??

You isolate a piece of genomic DNA that you think may be part of a natural product gene cluster. The DNA sequence of the fragment is attached as a text file on the 5.451 website (file “dnasequence”).

Go to the website <http://www.ncbi.nlm.nih.gov/>

On the left hand side of the screen click on Tools

Click on “ORF Finder” (about 5 paragraphs down the page)

Paste the DNA sequence into the box. Click on the button that says “orffind”

The computer has now found all possible ORFs (i.e. stop and start codons that are in frame) and highlighted them as turquoise boxes.

Click on one of the turquoise boxes. The program will translate the DNA into peptide sequence. At the top of the page click the button that says “BLAST”. BLAST is one of the most common algorithms for finding homologous proteins. Click the button that says “Format” on the screen that comes up. All homologous proteins will be listed, with the most similar proteins showing up first, and the least similar showing up last.

Find an ORF that encodes a gene involved in natural product biosynthesis.

Question:

The “unknown” DNA is most likely part of the biosynthetic pathway for what natural product??? (Click on the links of the homologous proteins to read about their function and thereby predict what natural product is encoded.)